DOCUMENT RESUME

ED 040 050 SE 008 293

AUTHOR Henkin, Leon; Kessner, Arthur
TITLE Mathematics 15.
INSTITUTION Cambridge Conference on School Mathematics, Newton, Mass.
SPONS AGENCY National Science Foundation, Washington, D.C.
PUB DATE 69
NOTE 227p.

EDRS PRICE EDRS Price MF-$1.00 HC Not Available from EDRS.
DESCRIPTORS *Curriculum Development, *Elementary School Teachers, *Instructional Materials, Mathematics, *Modern Mathematics, Number Systems, Teacher Education
IDENTIFIERS Cambridge on School Mathematics, National Science Foundation

ABSTRACT

Reports on the development of a course in modern mathematics for elementary school teachers. The subject matter for the course is concentrated on whole numbers, integers, and rational numbers. Ideas of logic and set theory are brought into the discussion incidently as needed for developing the number systems. Part 1 of the course is a brief essay on "What is a Teacher." Part 2 consists of four chapters on Number Systems. Each chapter is divided into sections, and each section is identified as belonging to one of three "tracks" which are interwoven throughout Number Systems. Track A presents basic mathematical ideas. Track B consists principally of ideas for work by means of which mastery of the conceptual material of Track A is gained. Finally, Track C presents some ideas to indicate how the conceptual material of Track A might be used by a teacher in the elementary school classroom. Part 3 is a sequence of notes on 16 lectures concerning an experiment in elementary school mathematics instruction. [Not available in hardcopy due to marginal legibility of original document.] (RP)

ED040050

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

MATHEMATICS 15

FALL 1969

by

Professor Leon Henkin
Department of Mathematics
University of California
Berkeley, California

Mr. Arthur Kessner
University of California
Berkeley, California

Cambridge Conference on School Mathematics

Funded by

Undergraduate Science Curriculum Improvement Division

National Science Foundation

SE 008 293

MATHEMATICS 15

Fall, 1969

Leon Henkin

and

Arthur Kessner

Prepared with the support of
Educational Development Center.

An Explanation

These notes consist of three parts. First there is a brief essay, WHAT IS A TEACHER. Then comes NUMBER SYSTEMS, consisting of four chapters. Finally there is a sequence of notes on lectures given in the Spring Quarter, 1968.

The lecture notes begin with page 4 of Lecture 11. The four chapters of NUMBER SYSTEMS are essentially an elaboration and reworking of the material which was contained in the first 11 lectures of the Spring, 1968 course. The notes were taken at my 1968 lectures by Arthur Kessner, a graduate student, and during the summer of 1968, he collaborated with me in the writing of Chapters 1-4 of NUMBER SYSTEMS. This work was supported by Educational Development Center.

The subject matter of the course is concentrated about three number systems: whole numbers, integers, and rational numbers. Ideas of logic and set theory are brought into the discussion incidently as needed for developing the number systems, and are not emphasized for their own sake.

The material in Chapters 1-4 of NUMBER SYSTEMS is organized in an unusual way. Each chapter is divided into sections, and each section is identified as belonging to one of three "tracks" which are interwoven throughout NUMBER SYSTEMS. Track A presents basic mathematical ideas. Track B consists principally of ideas for work by Mathematics 15 students, by means of which they can gain mastery of the conceptual material of Track A. Some of the work is in the form of exercises. Topics are suggested for possible discussion in the section meetings of the course. Occasionally a possible long-term project is suggested. Finally, in Track C we present some ideas to indicate how the conceptual material of Track A might be brought by a teacher into the elementary school classroom.

Berkeley
August, 1969 Leon Henkin

WHAT IS A TEACHER?*

Leon Henkin

Of course a teacher is someone who teaches -- not just occasionally, but someone who works at teaching. And teaching is helping to learn. But what is learning?

Most people will agree that learning is an activity you have to accomplish by yourself. It is like eating. Someone can tell you where food is, or can set it before you -- they can even put it in your mouth if you're a baby, say, or crippled. But the final act of eating you must perform yourself. And so with learning.

So helping someone to learn -- that is, teaching -- is a little like helping someone to eat. At the beginning a mother selects the food, buys it and brings it home, prepares it, puts it in the baby's mouth, wipes it off his chin, and puts it back in his mouth again. But her object is to get her child to eat independently. Ultimately he should be able to choose his own food to satisfy the requirements of both health and taste; he should be able to obtain, prepare, and eat his food himself.

And so, again, with learning. Unless a teacher helps her pupils to become independent of her, unless she consciously heads toward the day when they can do well by themselves, at choosing what to learn, at acquiring the necessary materials, and finally at learning, she will not succeed in the ultimate sense -- even if her pupils have gathered much information while they are with her.

If we take this viewpoint seriously, it has far-reaching implications for the organization of our schools. It does not mean that the teacher just gives heavy homework assingments. Rather, it means that assignments are designed to lead to genuinely independent thought, and that the activities in the school itself are directed toward encouraging students to pursue individual interests, to make discoveries, to acquire a taste for study, to develop an ability to gather information, and to understand.

What gives to teaching its greatest challenge, and what makes its problems so vastly more complex than those involved in helping a child to eat, is the tremendous variation in the learning process from one individual to another. The scientific study of learning is barely beginning, but it is recognized generally that there is a variety of basic patterns of learning, and that superimposed upon these patterns are the individual ability levels distributed over what are probably thousands of separate characteristics which enter into the learning apparatus of a given personality. Over and above the differences in predisposition and capacity to learn, the learning process, we know, is highly sensitive to the total experience of an individual -- his relations within the family and with friends, his contact with mass communication media, his reading, his dreams, his play. Our democratic aim of educating each person to the point where he can realize his fullest potential places upon the teacher a great responsibility; she should study and understand each one of her pupils as a distinctive individual, and devise ways of helping hem to learn what is peculiarly suited to his needs, his interest, and his ability.

---

* Reprinted from GOALS FOR MATHEMATICAL EDUCATION OF ELEMENTARY SCHOOL TEACHERS, A Report of the Cambridge Conference on Teacher Training, Houghton Mifflin Co., Boston, 1967.

Side by side with the continuing effort to analyze her students and to understand them sympathetically, a teacher has the obligation to work continuously at the selection of the facts, techniques, ideas, and attitudes which she will ask her students to learn, and at the development of the methods and devices/ she will employ to help them accomplish this learning. In this respect we must understand clearly that the nature of our society and the role of the individual in it are undergoing certain revolutionary changes and are disrupting patterns which have been constant heretofore for generations, if not for centuries; and we must clearly see that these changes impose upon the teacher a concomitant pattern of new duties. Nowhere is this clearer perhaps than in the area of mathematical instruction -- although in reality our developing ideas about physical and biological science, about the study of language (foreign and domestic), and the study of society, and our developing attitudes concerning intergroup relations, impose demands on the teacher which are just as heavy though perhaps less clearly articulated.

What are the changes in mathematics itself which must be reflected in the elementary classroom? For one thing, the sheer volume of new mathematical research has increased year by year at a sharply accelerating rate, and a significant fraction of this work affects our understanding of the most fundamental concepts. In direction, mathematics has become much more abstract and, paradoxically, because of this abstraction it has become applicable to, and has derived sustenance from, a much wider range of applications. From the study of numbers and geometric figures it has broadened its scope to include every domain where form and structure can be discerned. Finally, the art of computation has become infinitely more complex, and the practice of it has shifted the routine burdens of execution to electromechanical devices while demanding much more in the way of control and design from the practitioner.

All these developments require not only that the teacher must alter at this time the mathematical curriculum which has heretofore remained static but also that she must continue to alter it from year to year throughout her teaching career. They mean, too, that the teacher of elementary mathematics must work not merely at training students to follow and apply prescribed computational routines, but also at getting them to understand abstract concepts to the point where they can devise and test new computational routines; she must stimulate them to formulate new concepts arising from diverse realms of experience and to search for the properties which relate these to concepts; and she must educate them to employ relatively sophisticated patterns of mathematical language so that they can communicate freely about their work.

These multifold obligations which we are delineating for the elementary school teacher entail two principal positions about the individuals to whom we assign this work. In the first place, the amount of specialized knowledge, ability, and interest required to carry out these tasks in even a competent, not to say inspiring, manner, obviously transcends what can be expected of a single individual. We must begin to think of the elementary teaching corps as composed of a variety of persons contributing in diverse ways to a common goal. And, in the second place, we must recognize that the satisfactory discharge of her duties requires of a teacher many kinds of professional activity other than direct contact with students, and we must provide such working conditions and environment as will facilitate the prosecution of these activities. Let us examine these two propositions separately.

The idea of specialization among elementary teachers at first suggests the kind of instructional pattern now found at the high school levels. Most educators consider this pattern unsatisfactory at the elementary levels. Actually, however, the concept of specialized teaching is compatible with a host of instructional patterns quite different in character. For example, insofar as we require each child to be the subject of careful and sympathetic observation designed to discover optimum methods for developing his individual potential, it seems clear that each child should be associated with a single teacher, say "A", who can work with and study him in a variety of learning conditions and subject areas. At the same time, since "A" cannot be expected to be expert in all the everchanging curricular areas, we might wish a second teacher, "B", to divide her time among a larger group of children, seeing each child for only a short part of the day for the purpose of dealing specifically with a single subject area, such as mathematics. And of course we would expect "A" and "B" to work together, with respect to their common students, by intercommunication and joint projects. There thus emerges an instructional pattern very different from what we find in high schools and colleges, one in which a given pupil encounters more than one teacher simultaneously at certain parts of the day.

One implication of such a program is the prospect of a teaching force sharply increased in size. Such an increase may, in fact, prove necessary or desireable, but it is by no means required by a scheme of cooperative teaching. Of course, if our view of the teaching environment is restricted to the conventional, closed classroom containing thirty seats facing the front of the room, then simple arithmetic will show indeed that to have more than one teacher in a room at one time requires an increase in the total number of teachers. One the other hand, reverting to our early desideratum of independent, individual learning, we can imagine a school in which single students, or small groups, work by themselves for significant periods of time in semiprivate areas within open rooms of considerable size with several teachers, including "A" and "B", circulating and providing help and observation in a variety of ways. Such a program might well be feasible with no increase in total size of instructional staff.

It should be emphasized that there is wide latitude for experimentation in the cooperative teaching patterns which may be evolved to accomodate specialized teaching in the elementary school. The activities undertaken by elementary school teachers who have specialized in a subject area such as mathematics can range from dealing with individual students to conducting in-service courses for other teachers, to monitoring new text materials and innovations in technological teaching devices, to name but three examples; and we can envision groups of teachers each of whose activities are concentrated among a different few of the varied forms of specialization.

But let us look now at our second proposition: that an inventory of the new responsibilities being assigned to elementary school teachers requires us to provide adequate time and a suitable environment for a large variety of activities which do not involve direct contact with pupils. What does this mean in detail?

At present the normal distribution of an elementary school teacher's time involves seven hours a day at school: five are spent in direct supervision of students in class (or for brief periods on the playground), one is designated as a lunch hour(which often involves further contact with students in the cafeteria or even classroom), and one is spent in preparing the classroom before and after the students are in session and in a multitude of clerical tasks such as completing pupil attendance forms. Additional time in school is required on occasion to attend staff meetings, in-service training sessions, parent conferences, or PTA meetings. And work at home is expected to cover grading of student papers, preparation of lessons and inspection of text books. Essentially no time is provided for teachers to discuss problems involving curriculum or methodology; no time, space, or materials are provided for individual study, and the desirability of visits by teachers to pupils' homes is ignored. It is fairly evident that the most dedicated and efficient of teachers can barely be expected to discharge even the most traditional instructional duties in a satisfactory manner within such painful circumstances. To expect her to measure up, under such conditions, to the kind of job we have outlined is absurd.

What sort of school environment can we imagine within which we could realistically expect a teacher to perform the tasks we have listed? Perhaps an average of three hours per day of contact with the pupils is a reasonable goal at which to aim. Subprofessional teacher's aides must be engaged to provide clerical assistance and to free the teacher from routine chores so that she can concentrate on her primary work. Help with homework and playground supervision is desireable, though the teacher will wish to maintain some contact with these student activities. Space and facilities for individual study, or for conferences among small groups of teachers, must be available; the most basic library items, such as current journals in education and subject areas, should be on hand in a suitable room of the school; and time for access to more substantial library centers in the school district should be considered a normal part of the workday. Both subject specialists and general teachers will wish to study their individual students in depth, and each school indubitably will have ably manned facilities for psychological testing and research. Those aspects of instruction which can be made routine, such as drill in the application of algorithms, will be assigned to computer-based machines, so that the time of teachers may be conserved for the truly creative aspects of instruction. Films and projection rooms, as well as properly equipped laboratories for science, language study, and mathematics, will be among the teacher's tools. Administrative assistance and computerized aids will facilitate the testing and regrouping of students where instruction is organized in ungraded patterns. The teacher will be expected and encouraged to continue her studies in the material of her subject area as well as in methodology, and a regular system of sabbatical leaves will be avail- for those teachers wishing to deepen their penetration of some aspect of their work.

It will not have escaped the reader's attention that the creation of such a stimulating environment for teaching and learning is likely to involve considerable increase in the cost of education. To some extent the economic cost may be lightened because the assignment of routine instructional tasks to computer-based machines and the widespread and persistent encouragement of individualized self-study will tend to lower the teacher-pupil ratio. In this connection we may expect to see increasing use of students themselves as instructional aides, for learning-by-teaching is a phenomenon long familiar to teachers but not adequately exploited as a systematic method of instruction.

Any savings effected by reducing the teacher-student ratio will be more than offset, however, by the increased salaries which will have to be provided to assure an adequate supply of sufficiently talented individuals. The present miserable salary scales prevailing in most school districts take improper advantage of the fact that the great bulk of elementary teachers do not need to support a family with their income from teaching. But for the teaching profession of the future we shall have to attract many men and women who can fashion a full-time, lifelong career upon their work. Indeed, for this purpose we shall have to do much more than raise salaries. We shall have to create a structure within the profession of elementary teaching whereby merit is given due recognition, and wherein advancement from one level of specialized teaching to another, involving increased responsibility and rewards, can continue over a long span of years without being diverted from teaching to administration.

And so we return to the question with whch we began -- What is a teacher? In part the answer will be determined by the decisions of society, by the demands which it places upon teachers, and the support it provides. In part the answer will be determined by the decisions of of these individuals who dedicate themselves to a career of scholarship and of "helping to learn." It is not overly dramatic to say that these decisions will turn out to be critical in shaping the destiny of mankind.

It falls to us, the contemporary scholars and educators, to point out the possibilities to society and to inspire the individuals who will follow us.

NUMBER SYSTEMS

Notes for Prospective Elementary Teachers

by

Leon Henkin
and
Arthur Kessner

Prepared with the support of Educational Development Center

Part 1: The system of whole numbers

Chapter 1: Number systems; what we learn first.

§1 (Track A)

A number system consists of a set of objects called numbers, together with certain things we call operations on the numbers. In Part 1 we shall study the system of whole numbers, 0, 1, 2, 3, ... , together with the familiar operations of addition (+) and multiplication (·) on these numbers. In Part 2 we shall study number systems involving the integers, which in addition to the whole numbers include the negative integers -1, -2, -3, ..., and in Part 3 we shall deal with number systems involving the rational numbers, which in addition to the integers include numbers such as $\frac{5}{3}$, - 12/7, etc.

There are too many whole numbers for us to write down a name for every one of them, so we often content ourselves with writing names for the first few and then using three dots to indicate the others: 0, 1, 2, ... We shall use the letter "W" as a name for the set whose elements are all of the whole numbers 0, 1, 2, ...

If we have written the names for several objects, say the Eiffel Tower and the city of Moscow, and if we wish to talk about the set having those objects as its elements, we form a name for this set by enclosing the list of objects within braces. For example, {Eiffel

Tower, Moscow} is the set whose two elements are the indicated tower and city. Using this convention we see that {0, 1, 2, ...} is the set of all whole numbers, that is, the set W. We express this by the sentence

$$\{0, 1, 2, \ldots\} = W.$$

The equality sign means that the thing named (or described) on the left is the very same object as the thing named on the right. (For example, we can write: Pierre Curie = the husband of Marie Curie.)

If A is a set having several elements (of any sort whatever), we can attach a whole number to A by a process called counting its elements. This process is usually the way in which numbers are first introduced to young children. We shall use the symbol "n(A)" to denote the number obtained by counting the elements of A; in other words, n(A) = the number of elements in A. For example,

$$n(\{\text{Moscow, Eiffel Tower}\}) = 2.$$

A child learns to count by imitation. In order for us to understand the process of counting we must analyze the components which enter into this process.

In the first place, an obviously important component is the recital of the numbers in a certain fixed order: one, two, three, ... If a child begins mixing up the order of the numbers, say one, four, five, two, ..., he is not likely to arrive at a correct count.

In the second place, the numbers thus recited in their standard order are "attached" to the objects being counted -- for example, by touching the objects one at a time as the numbers are called out. This is often a difficult process, for every object must be touched (i.e., must have a number attached to it), yet no object may be touched more than once. If we continue this process we arrive at a last element of the set A, that is, an element such that, after it has been touched, there are no elements of A left untouched.

The number called out as we touch this last element of the set A is n(A) -- the number of elements in A.

If there are very few elements in A, it is not too hard to remember which elements have already been touched (i.e., counted) at each stage in the counting process. If, however, there are a large number of elements in A, if they happen to look very like one another, and especially if they are scattered in a very disorderly fashion, then the process of counting by touching can become exceedingly difficult -- for a grown person as well as for a child. For under these conditions it becomes difficult to tell, after a while, whether a given element of A has been left out or has been counted before.

What are some of the ways we can ease the burden of a difficult counting job? One way is to relieve our memory by marking each object in some way as we touch it; even so, in a great crowd of objects we may not find it easy to be sure whether, at a certain stage, every object has been marked. A quite different method is to "line up" the elements of A (if, that is, they can be easily moved about). Once the elements have been arranged in a linear array, one after the other, the counting process becomes greatly simplified.

Indeed, if the objects of a set are all lined up in a row, about the only thing that can go wrong when a child tries to count them is if he runs out of numbers. That is, the child touches one of the objects and calls out a number, then he moves on to the next object but he doesn't know which number comes after the last one he called out.

Of course this doesn't often happen to grown-ups in our society, but among young children it happens frequently. First a child may learn to count to three; then to ten, then to thirty, then to ninety-nine. Eventually, he no longer takes pride in how far he can count because he comes to understand the systematic method of naming the numbers in their standard order (that is, the numeration scheme). He then sees that he can count "as far as he wants".

§2 (Track B)

1. Notation for sets. A set is an abstract object associated with certain other objects which are called its elements, or members. Given any objects whatever, there is a set having no other elements but them. Furthermore, the elements completely determine a set -- it is impossible to

have two different sets which have exactly the same elements. (This is known as the principle of extensionality.) For this reason {0, 1, 2, 3}, the set whose elements are the numbers 0, 1, 2, and 3, is the same set as {0, 3, 1, 2}, the set whose elements are the numbers 0, 3, 1, and 2. That is, we have {0, 1, 2, 3} = {0, 3, 1, 2}. We see from the principle of extensionality that if A and B are two different sets, then one of them must have an element which is not an element of the other.

The set {0, 2, 0, 3, 0, 1, 4, 3} is the set whose elements are the number 0 and the number 2 and the number 0 and the number 3 and the number 0 and the number 1 and the number 4 and the number 3; that is, it is the set whose elements are the numbers 0 and 2 and 3 and 1 and 4. Hence we may write {0, 1, 2, 3, 4} = {0, 2, 0, 3, 0, 1, 4, 3}. Furthermore, we have {0, 1, 2, 3, 4} = the set of all whole numbers less than 5.

2. Exercise. Consider the following sets Z, H, B, C, and D.

Z = {0, 2, 4, 8, 6}
H = the set of all female presidents of the U.S.A. during the 19th century.
B = {8, 4, 4, 0, 2, 0, 6}
C = { }
D = the set of all even whole numbers less than 8.

(a) For each pair of distinct letters chosen from among "Z", "H", "B", "C", and "D" form a true statement by employing one of the symbols " = " or " ≠ " between the letters.

(b) How many elements does the set {Z, H, B, C, D} have?

(c) n(Z) = 5. Compute n(H), n(B), n(C), and n(D).

3. Empty sets. The set H (in the previous exercise) is said to be empty because it has no elements. Similarly, the set C is empty. Actually, there is only one empty set, since if A and B are any sets each of which is empty, then we must have A = B, because if we had A ≠ B then either A or B would have to contain an element which is not in the other -- by the principle of extensionality mentioned in item 1 above. A common name for the empty set is " ϕ ". Thus ϕ = { }.

4. Exercise. Give three different descriptions of the empty set. (Compare the description of the set H in Exercise 2 above.)

5. Exercise. Indicate whether each of the following sentences is true or false. Justify your answer.

(a) If x,y are any persons such that x = y, then also the father of x = the father of y.

(b) If x,y are any persons such that the father of x = the father of y, then also x = y.

6. Classroom discussion. Discuss ways of determining whether two given sets have the same number of elements, and if not which set has more elements. Some ways do not require the use of numbers at all. Are there ways which apply to sets of movable objects which cannot be used with two sets of planted trees?

7. Classroom discussion. Using only the 3 letters "A", "B", and "C", discuss several different methods for combining them to obtain a system of names for each of the whole numbers 0, 1, ..., 20. Recall the Roman Numeral System -- how many whole numbers are named in this system using only the letters "I", "V", and "X"?

§3 (Track C)

1. Describing objects and sets. At the kindergarten and first-grade levels there is an abundant supply of small physical objects in every classroom -- crayons, blocks, books, for example. Give the children much practice in describing sets having these objects as elements. At first they will form "naturally grouped" elements, such as all the crayons in a certain box, or all the books on a certain shelf. But they should then be led to form other sets consisting, for example, of a certain book and two certain crayons. At first the specific items may be identified by pointing and saying "This", but children should be led to

describe objects without pointing, such as "The paper on the teacher's desk which is nearest the window". Make a game out of this.

Children should be led to bestow letters as temporary names for designated objects, e.g., "Let A be the window at the front of the room". Then they should use the symbols " = " and " ≠ " between such letters to form true statements. (They can thus write sentences before they can spell.) They should also use the phrases "equals" and "is unequal to" orally, as in "The piece of chalk on the floor equals the piece of chalk used by Robert to write his name".

Introduction of the empty set will provide lots of fun as children delight in impossibilities: Each child should give some description of the empty set, e.g., "the set of children with 5 ears", or "the set of chairs hanging from the ceiling".

Describing unit sets, i.e., sets containing exactly one element, can also be fun. Is the set of all boys in the class who have a sister named Sue, a unit set? If not, how about the set of all brown-eyed boys in the class who have a sister named Sue? Have each child describe a set of people in the classroom, and discuss the question whether there are two different descriptions of the same set; use the language "equals" or "is unequal to" between two set-descriptions, as appropriate.

§ 4 (Track A)

The need for a schematic way of naming the whole numbers arises because the numbers go on and on -- there is no last one. However far along we get in counting, there is always a next number beyond the one we are at. For this reason we can never hope to give individual names to the successive numbers in an unrelated way, as we can to the successive children of a family. If we want each whole number to have a distinctive name, we need a numeration scheme which tells how, given the name of any number, we can form the name of the next number from it.

As we have indicated above, at a certain stage children "catch on" to this scheme, and then they lose interest in the game of counting higher and higher. But the scheme is seldom described explicitly to them by a teacher, and indeed a clear and full mathematical description of the numeration scheme becomes quite difficult and sophisticated. Let us see how one would begin.

First, of course, come several numbers to which we do give separate and independent names. These are the first ten whole numbers, denoted by the familiar arabic numerals: 0, 1, 2, 3, 4, 5, 6, 7, 8, 9. The order of these is established in the manner indicated by the display in the previous sentence, and this plays a key role in describing the scheme for naming the later numbers. In order to have a convenient way of referring to this order let us agree that whenever we are given any whole number x, the number which comes next after x will be called *the successor of x*, and will be denoted $S(x)$ for short. Thus $S(0) = 1$, $S(1) = 2$, $S(2) = 3$, etc.

Now then, after the numbers 0, 1, ..., 9 comes another set of whole numbers each of which will be named by a *pair* of the arabic numerals, written one after the other. The right numeral of this pair may be any of the ten numerals whatever; the left one may be any numeral *except 0*. Such a pair we will call a *two-digit numeral*.

Now in order to describe which two-digit numeral will denote which whole number, we have to specify (a) which two-digit numeral comes first, i.e., which one denotes $S(9)$, the next whole number after 9, and (b) a rule which tells us how, given any two-digit numeral xy which denotes a whole number, we obtain the two-digit numeral which denotes successor of that number. This we do as follows:

Rule for using the two-digit numerals.

(a) $S(9)$ is denoted by the two-digit numeral 10;

(b) Given any two-digit numeral xy, the successor of the whole number denoted by xy is denoted by $x\,S(y)$ if y is not 9 (case 1), or by $S(x)\,0$ if y is 9 and x is not 9 (case 2). In case both x and y are 9, then the successor of the number denoted by xy is not denoted by any two-digit numeral, but is the first of a series of numbers which are denoted by three-digit numerals.

It is probably worthwhile for the reader to elaborate the ideas used in the preceding paragraphs to give a precise mathematical description of the scheme for using three-digit numerals as names for certain whole numbers which follow those denoted by the two-digit numerals. He will then see how, in principle, one could describe the scheme for using numerals composed of any fixed number of digits. However, a single description of the scheme covering the use of numerals containing an arbitrary number of digits requires the use of a somewhat sophisticated principle called mathematical induction, which we shall describe in a later chapter.

As we have seen, the process of counting, in order to obtain the number n(A) which indicates how many elements are in a given set A, involves a fixed ordering of the whole numbers and the attachment of these numbers, in order, to the elements of the set A. The process of attaching numbers to elements may be touching-while-reciting, or it could consist of making tags on which the numerals (i.e., the names of the numbers) are written and then tying these tags to the elements of A, or it could consist in writing names of the elements of A in a list opposite the names of the numbers, or it could be by still other methods. From the mathematical point of view it makes no difference which method of attaching numbers to the elements of A is employed. The only thing that matters is to know which number is attached to which object, not how the attachment is brought about.

In order to deal with the question of attaching numbers to objects in a neutral way which can represent any possible manner of causing the attachment, the mathematician has invented the notion of a function. A function can be thought of as an abstract device for indicating which number is attached to which object in a given set. If the function is denoted by the letter "f" and if "b" denotes some element of the set A, then we use the notation "f(b)" to indicate the number attached to the element b by the function f.

For example, if we wish to count the elements of the set {Moscow, Eiffel Tower}, we could do so by means of the function f such that

$$f\ (\text{Moscow}) = 1 \quad \text{and}$$
$$f\ (\text{Eiffel Tower}) = 2.$$

We get the same function f whether we paint the numerals "1" and "2" somewhere on the city and the tower, or whether we write the numerals on balloons and float them over the city and the tower -- just so long as it is the "1" which is attached to the city and the "2" to the tower. Another way to count the elements of the same set would be by means of the function g such that

$$g\ (\text{Moscow}) = 2 \quad \text{and}$$
$$g\ (\text{Eiffel Tower}) = 1.$$

Using the function f, the last element to get counted is the Eiffel Tower; using the function g, the last element to get counted is the city of Moscow. Either way, the number which is attached to the last element is 2. Hence

$$n(\{\text{Moscow, Eiffel Tower}\}) = 2.$$

For the particular set just considered it is easy to see that the specified functions f and g give the only possible orders in which the elements of the set can be counted, and of course in both cases the same number gets attached to the last element to be counted. But will a similar result be true in counting the elements of any set whatever? Suppose, for example, we had a very large set, say B, and that one way of counting its elements ended by attaching a certain whole number x to the last element to be counted, while a different order of counting the elements of B ended with a number y, different from x, being attached to the last element to be counted. Then what would n(B) be? Would we have $n(B) = x$ or $n(B) = y$? Would the phrase "the number of elements in B" have any sense at all?

The fact is that this situation cannot occur. The whole concept of "the number of elements in the set B" is based on the supposition that there is a unique number, n(B), which will be arrived at by any two correct ways of counting the elements of B. But how do we know that no matter what set B we start with, any two methods of counting its elements will end up by attaching the same number to the last counted element? As children, we come to believe this fact after trying out the counting process on a few sets and then developing an intuitive conviction that the same result will come about for other sets which we have not tried to count. As mathematicians, however, we may not be satisfied to base the whole theory of counting on such an intuitive guess: We may seek somehow to prove the result that any two methods of counting the elements of a given set must give the same result. Such a proof can be carried out in a branch of mathematics known as theory of sets, but we shall not do so here.

Can every set be counted? Or are there sets which in some sense have too many elements to be counted? The set W of all whole numbers is an example of a set which we say is infinite, or has infinitely many elements: There is no whole number n(W) which tells how many elements are in W.

The way we see this is in noting that when we line up all the whole numbers, say in their natural order, there is no last one. (As noted previously, every whole number has a successor.) But the whole process of counting the elements of a set depends on arriving at a last element and seeing what number is attached to it. If there is no last element, then the process of counting does not terminate and so does not give a result.

We may conclude from the above discussion that an infinite set is one whose elements can be lined up in such a way that there is no last one. However, there is one exception to this rule. What about a set which has no elements?

Can there be such a set?

We have, indeed, encountered empty sets in item 3 of §2 above. Mathematicians find it convenient for many reasons to deal with a set having no elements -- so they invent one! As indicated in §2 the empty set is often denoted by the symbol "ϕ". Of course we cannot line up the elements of ϕ so as to obtain a last one; yet we do not wish to call ϕ an infinite set. We say that the number of elements in ϕ is 0, and write

$$n(\phi) = 0.$$

This is why, when we begin counting the elements of a set which is not empty, we begin with the number 1. The numbers 1, 2, 3, ... are sometimes called the counting numbers. Mathematicians also call them positive integers.

§5 (Track B)

1. Exercise. The set W of all whole numbers is an example of an infinite set. Give an example of a different infinite set, all of whose elements are whole numbers. Also give an example of an infinite set having some elements which are not whole numbers, and an example of an infinite set none of whose elements is a whole number. In each case justify the statement that the set in question is infinite.

2. Exercise. A set which is not infinite is called finite. Let G be the set whose elements are all finite

sets of whole numbers. Is G finite or infinite? Justify your answer.

3. Exercise. Recall that if x is any whole number, then S(x) is its successor, the next whole number after x in the natural (counting) order. Now, suppose that x and y are whole numbers such that

$$S(S(S(x))) = S(S(y)).$$

Do we necessarily have $y = S(x)$? Justify your answer.

4. Exercise. In a later chapter we shall study the operation of addition, + . Using your intuitive knowledge of this operation, determine in each case below a number y which satisfies the indicated condition.

(a) $S(S(y)) = 6$

(b) $S(4+2) = 4 + S(y)$

(c) $3 + 2 = S(S(S(y)))$

(d) $y + 4 = S(S(S(S(S(4)))))$

(e) $5 + y = S(S(S(2)))$.

A number y which satisfies an equation (i.e., makes the equation true), is called a solution of the equation. The set of all solutions to a given equation is called the solution set of the equation. Make up an equation similar to those above whose solution set is empty.

5. *Exercise*. (a) Employ the *Rule for using the two-digit numerals*, given in §4, to determine the second two-digit numeral and the fourth two-digit numeral, explaining how you obtained these by the rule.

(b) Define what the *three-digit numerals* are, and make up a Rule for using three-digit numerals, following the general pattern used in §4 for two-digit numerals and modifying that pattern as seems appropriate. (Hint: What is the first three-digit numeral, i.e., the three-digit numeral which denotes the successor of the number denoted by the last two-digit numeral? Then describe how, given an *arbitrary* three-digit numeral, we obtain the next one; this description will involve enumerating several cases.)

6. *Exercise*. Suppose a child has three checkers which we shall call a, b, and c. If he counts them by touching them in the order b, c, a as he calls out "one, two, three" he is attaching the number 3 to a, 1 to b, and 2 to c. Mathematically speaking we say that he is employing the counting function f such that

$$f(a) = 3, \ f(b) = 1, \text{ and } f(c) = 2.$$

We could also describe f by giving a table of values for it, as follows:

x	f(x)
a	3
b	1
c	2

(i) Give equations describing the counting function g employed when counting the checkers in the order c, a, b.

(ii) How many different counting functions are there altogether for the set of checkers {a, b, c}? Make a table showing the values for all of these counting functions.

(iii) How many counting functions are there for a set of 5 marbles?

(iv) Try to generalize the results of (ii) and (iii) above by finding a rule which tells how to compute the number of counting functions for a set of n objects, where n may be any whole number whatever.

7. Exercise. Think of the numerals 0, 1, ..., 9 as an alphabet, and the numerals 0, 1, ..., 20 as words of one and two letters made up from this alphabet. Arrange these words in "alphabetical order".

Imagine all numerals (with any number of places) arranged in this "alphabetical order". Given any numeral n, let $S_a(n)$ be the next numeral after n in this order. What is the first numeral n after o such that $n \neq S_a(p)$ for every numeral p?

Chapter 2. Successor and order

§1 (Track A)

In the previous chapter we considered the set W of all whole numbers, 0, 1, 2, ..., and we discussed the way in which these numbers are used to indicate how many elements are in a given set A. The number of elements in the set, n(A), is determined by a process of counting, and we mentioned certain mathematical concepts which underlie the counting process, namely, the standard _ordering_ of the whole numbers based upon the fact that each of these numbers has a _successor_, and the method of attaching numbers to the elements of a set as embodied in the concept of a _function_.

In the elementary schools it is customary to follow the study of counting by introducing the operation of addition, and we shall adopt a similar course here. However, before proceeding to study addition we wish to examine in more detail those mathematical concepts we have found to underlie the counting process.

Let us start with the notion of _successor_. We have agreed that if x is any element of W, that is, any whole number, then it has a _successor_, which is the next number following x in the natural order. We denote the successor of x by the notation S(x). Now what about the symbol S itself -- does it have a separate meaning? The mathematician considers that this symbol is the name of an _operation_:

S is an abstract operation which can act on any whole number, that is, on any element of W, and the result of this action is another element of W. For example, when the operation S acts on the number 14 it produces 15, and we express this by writing S(14) = 15. Similarly, S(1039) = 1040. We say that S is the successor operation.

Of course the word "operation" is familiar to us from elementary-school mathematics, for we speak about the operation of addition, or the operation of multiplication. In contrast to the successor operation, these operations act on pairs of numbers, instead of on single numbers. For example, when the operation of addition, +, acts on a pair of numbers (x,y), we denote the resulting number by the notation x + y; in particular we have 2 + 35 = 37, which indicates that when the operation + acts on the pair (2,35), the number which results is 37.

We distinguish operations like S from operations like + and · by saying that the former is a one-place operation on W, while the latter are two-place operations on W.

Clearly we can expect one-place operations to be simpler things to study than two-place operations. Since the operation S plays such a basic role in the system of whole numbers, it is surprising, therefore, that it is not given a name and studied explicitly in the elementary school. Let us investigate its mathematical properties.

First of all, because we have a numeration scheme which provides a name for every one of the infinitely many whole numbers and which tells us how to find the name of the successor of any given number, we can write down infinitely many true sentences about S, such as

$$S(0) = 1, \quad S(28) = 29, \quad S(1{,}386) = 1{,}387, \quad \text{etc.}$$

These are examples of particular statements involving the successor operation, S, for each statement involves two particular specified numbers. Particular statements of this kind are often summarized in a table, as follows:

x	S(x)
0	1
28	29
1,386	1,387
⋮	⋮

Such a table is called a table of values of the operation S.

There are other particular statements about S which are true, besides those which are summarized in the table. One of these, for example, is

(a) $S(S(3)) = 5$.

It is useful for us to notice that this fact can be inferred, by a process of logical deduction, from the following two statements (which do express entries from the table of values for S):

(b) $S(3) = 4$, and

(c) $S(4) = 5$.

How would we carry out such a deduction, or proof?

First, starting from statement (b), we would conclude

(d) $S(S(3)) = S(4)$,

by the logical meaning of the symbol = . The point is that statement (b) assures us that $S(3)$ and 4 are the same number; hence if we perform any operation on $S(3)$ we of course get the same result as of we perform that operation on 4. In particular, $S(S(3))$ is the same thing as $S(4)$, which is the content of (d).

The next stage of the proof is to combine equations (d) and (c) to get the desired result (a). This step is again justified by the logical meaning of the symbol = , for (d) tells us that $S(S(3))$ is the same number as $S(4)$, (c) tells us that $S(4)$ is the same number as 5, and so of course $S(S(3))$ is the same as 5 -- which is the content of (c).

The whole deduction described above would normally be summarized in the following form.

Theorem. From the hypotheses $S(3) = 4$ and $S(4) = 5$, we may infer $S(S(3)) = 5$.

Proof.

(1) $S(3) = 4$; hypothesis.

(2) $S(S(3)) = S(4)$; logic of = .

(3) $S(4) = 5$; hypothesis.

(4) $S(S(3)) = 5$; by logic of = from lines (2),(3).

Statements of the form $S(S(1)) = S(2)$ are called equations; they assert that a number formed in one way (in this

$$S(0) \neq 0, \quad S(1) \neq 0, \quad S(2) \neq 0, \quad \ldots .$$

Another important general property of the successor operation is the fact that whenever it operates on two different numbers it gives two different results. We express this by writing:

Proposition 2. For any whole numbers x,y such that $x \neq y$ we have also $S(x) \neq S(y)$.

Mathematicians have a name for this property of S: We say that the operation S is one-one (read "one-to-one"). Another way of writing the same idea is this: If x,y are any whole numbers such that $S(x) = S(y)$, we have also $x = y$.

Other general statements about the successor function are: For all x in W, $x \neq S(x)$; for all x in W, $x \neq S(S(x))$; etc.

§3 (Track B)

1. The identity operation. Let us introduce another operation, even simpler than S, of whole numbers: the one-place operation of Identity, which is written I. It is the "nothing-happens" operation, because when I operates on any particular whole number the result is simply that same whole number itself. For example, $I(0) = 0$, $I(1) = 1$, $I(49) = 49$.

2. Exercise. (a) Make a table of values for the opertion I, showing values of I for three or four numbers. Also, give a few examples of particular statements which are true of the identity operation, I, and involve the diversity

symbol, $\neq$. Finally, give several true particular statements which involve repeated use of the operation I.

(b) From the hypotheses $I(99) = 99$ and the logic of equality, deduce the particular statement $I(I(I(99))) = 99$.

(c) Express the following general statement about the operation I using letters as variables: When I operates on any whole number, the result is that same whole number.

(d) Use variables to express the true general statement that the operation I is one-one. (Hint: refer to the text's discussion of the one-one property of the operation S. See Proposition 2, §2.)

(e) Use the diversity symbol, $\neq$, to express two general statements involving the operation I.

(f) Consider the statement: If x is any whole number, then $I(x) \neq S(x)$. Is this a particular statement or a general statement? Is it true or false? Justify your answers.

3. Exercise. Using propositions 1 and 2 of §2, prove the general statement: If x is any whole number then $S(S(x)) \neq S(0)$.

4. Classroom discussion. Have the class review orally all that they know about the operation S, calling specific attention to the meaning of the new terminology involved with S, such as "operation", "one-place", "one-one", etc.

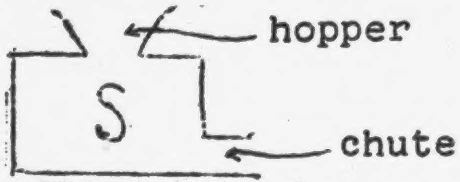
5. Classroom discussion. Consider the two equivalent forms of Proposition 2, §2. With a minimum of logical

terminology discuss the intuitive basis for considering these two sentences equivalent.

Proposition 2 expresses the fact that the whole number 0 is not the successor of any whole number. Actually, 0 is the only whole number with this property. One way to express this fact, using variables, is as follows: If y is any element of W such that $S(x) \neq y$ for every whole number x, then $y = 0$. Students should find an equivalent way of expressing the idea, beginning: If y is any element of W which is different from 0, then ... .

§ 4 (Track C)

1. The phrase "successor of a whole number" can be introduced successfully in the earliest grades as a synonym for "the next whole number". One way to generate interest in and get practice with the successor operation would be to picture S as a kind of machine, such as pictured here.



A whole number, say 4, is put into the hopper and out of the chute at the bottom comes its successor, 5. Let the children run the machine -- that is, have a child go to the blackboard where the machine is pictured and have him ask others for numbers to put into the hopper. He then tells what number comes out of the chute.

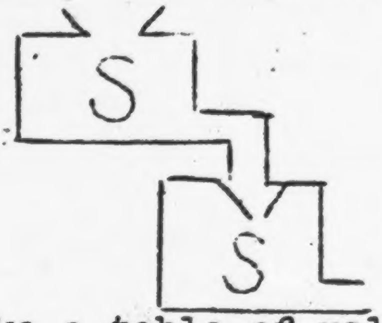
Later the procedure can be altered so that the children are asked to give the number that went into the hopper after being told the number that appeared at the chute. For example, given that 6 appears at the chute, what number went into the hopper? (In this case 5). Notice, however, that with this machine 0 could not appear at the chute, whereas any whole number can go into the hopper.

After playing with this machine for a while the pupils can be introduced to the idea of a table of values for the "successor machine". By setting down the results of "observation" in a systematic manner:

put in	came out
3	4
8	9
2	3

They are led to a method of recording findings which will lay the groundwork for later work with addition and multiplication tables.

After working with single machines for a while, children can experiment with hooking machines together, so that the output of one is fed directly into another.



The children should make a table of values for this double-machine, feeding into its hopper each of the numbers 0, ..., 9. This should be compared with the table of values of a single

S-machine with the same input. What relationship is observed between the output columns of the two tables? Can the children explain this?

§5 (Track A)

So far we have dealt with the successor function only as an abstract operation on whole numbers. However, we know that these numbers are applied to counting the elements of a given set. It is natural to inquire what the relation of the operation S is to the counting process. It turns out that there is a simple answer. Namely, if x is the number of elements in a set A, then S(x) is the number of elements of any set obtained from A by putting in one new element. Let us introduce some mathematical notation for referring to such a set.

Suppose, then, that A is any set, and suppose that b is any object which is not an element of A. Let us first form the set $\{b\}$ having b as its only element, and then let us put together the elements of the two sets A and $\{b\}$ to form one big set; we call this the union of A and $\{b\}$, and use the notation $A \cup \{b\}$ to denote it. Thus $A \cup \{b\}$ is the set formed from A by adding the object b as a new element. Now, how many elements are there in this new set, $A \cup \{b\}$? As we have indicated above, it is pretty obvious that the number of elements in $A \cup \{b\}$ is the next number after n(A), i.e., we have $n(A \cup \{b\}) = S(n(A))$.

Let us formulate this observation as a proposition for later reference.

Proposition 1. If A is any set and b is any object which is not an element of A, then $n(A \cup \{b\}) = S(n(A))$.

For example, if A = {Cairo, Jerusalem} and b = Suez Canal, then $n(A) = 2$, $S(n(A)) = S(2) = 3$, $n(A \cup \{b\}) = n(\{\text{Cairo, Jerusalem, Suez Canal}\}) = 3$.

Although the result expressed in Proposition 1 is intuitively very clear, it is worthwhile to explain it in terms of the counting process which underlies the determination of $n(A \cup \{b\})$. If we count the elements of the set $A \cup \{b\}$, we may first count those elements which are in A. If $n(A) = x$, the number x will be attached to the last element of A to be counted. At that point there will remain one uncounted element of $A \cup \{b\}$, namely, the object b. The rules of counting prescribe that the number attached to this element must be the next number after the number x just used. This next number is, of course, $S(x)$, which thus gets attached to b. Since b is the last element of $A \cup \{b\}$ to be counted, this gives the desired result, $n(A \cup \{b\}) = S(x)$.

§ 6 (Track B)

1. Union of sets. The two-place operation on sets called union was introduced in §5 as a method for taking any two given sets and forming a new one from them by combining all their elements into one big set. For example, if $A = \{1, 5, 14, 16\}$ and $B = \{7, 6, 5, 1, 16\}$ we get immediately that the union, $A \cup B$, is the set $\{1, 5, 14, 16, 7, 6, 5, 1, 16\}$.

Of course, repeating the name of an element, such as 1 or 5 in this case, does not in any way change the fact that 5 is an element of $A \cup B$ in exactly the same way as 7 is an element of this set. Similarly, in listing the elements of a set the order in which the list is presented is immaterial (as we see by the principle of extensionality). Hence we can also write $A \cup B = \{1, 5, 6, 7, 16, 14\}$.

2. Exercise. a) Why is the operation union called a two-place operation on sets?

b) Let $C = \{3, 6, 4, 16\}$, $B = \{18, \text{Moscow}, 6, 3\}$, $G = \{16, 6\}$, and $F = \emptyset$.

(i) List the elements of the sets $C \cup B$, $C \cup G$, $F \cup G$, $B \cup G$.

(ii) Determine $n(C)$, $n(B)$, $n(G)$, $n(F)$, $n(C \cup B)$, $n(C \cup G)$, $n(F \cup E)$ and $n(B \cup E)$.

(iii) Notice that $n(F \cup G) = n(F) + n(G)$. Is this relation true in general, that is for any sets F and G? Justification?

c) If $E = \{0, 2, 4, 6, 8, \ldots\}$ and D = the set of all odd whole numbers,

(i) What is $E \cup D$?

(ii) Notice that $E \cup D = D \cup E$. Is this true for other sets E and D?

d) Let A be any sets. How are the sets $A \cup A$ and $A \cup \emptyset$ related to A?

e) Let A be any set. We know that if b is any object which is _not_ an element of A, then $A \cup \{b\}$ is the set obtained from A by adding b as a new element, and $n(A \cup \{b\}) = S(n(A))$. But suppose, now, that b _is_ an element of A. How is the set $A \cup \{b\}$ related to A in _this_ case? And how is $n(A \cup \{b\})$ related to $n(A)$ in this case?

3. _Discussion_. Discuss the meaning of some of the notation introduced in this chapter. For example, which of the expressions "$S(n(A))$" and "$n(S(A))$" is meaningful -- where "A" denotes a set -- and what is its meaning? If "x" denotes a whole number, is "$n(S(x))$" meaningful? What about "$n(\{S(x)\})$"?

4. _Counting functions_. Recall that the mathematical notion of a _function_ may be used to indicate which number is attached to each element of a given set A in the process of _counting_ the elements of A. The numbers which get attached to the various elements of A in the counting process are the numbers 1, 2, ..., $n(A)$; the last number, $n(A)$, which is attached to an element of A in the counting process, tells us how many elements the set A has. A function f which attaches the numbers 1, 2, ..., $n(A)$ to the elements of A in a one-to-one manner is called a _counting function_. Mathematicians also study different functions which may attach objects other than numbers to the elements of a given set A, or which may attach the same number to several different elements of A.

5. _Exercise_. Let $A = \{0, 2\}$ and $B = \{0, 2, 5\}$. There are just two counting functions for A, namely, the function f such that $f(0) = 1$ and $f(2) = 2$, and the function g such that

$g(0) = 2$ and $g(2) = 1$. How many counting functions are there for B?

6. Exercise. Suppose A is any (finite) set, b is not an element of A, and $C = A \cup \{b\}$. Let x be the number of counting functions for A, and let y be the number of counting functions for C. Can you find a simple equation connecting the numbers x, y, and n(C)?

7. Let $A = \{0, 2, 4, 6\}$ and B = the set of all whole numbers less than 12.

a) Find a set C such that $A \cup C = B$. How many different sets C of this kind are there? How many of these are disjoint from A?

b) Find a set D such that $A \cup D = A$. How many different sets D of this kind are there? How many are disjoint from A?

§ 7 (Track A)

Let us now turn to the concept of order for the set W of all whole numbers. We have seen that all of the elements of W can be produced by starting with 0 and successively applying the operation S. Of any two distinct whole numbers, say x and y, one will be produced before the other by this process. If x, say, is produced before y, we say that x is less than y, or x is smaller than y, and we write $x < y$. For example, we have $2 < 14$, $23 < 678$, but not $3 < 0$.

The fact that of any two distinct whole numbers one must be less than the other, can be formulated as a general statement:

For all whole numbers x and y such that $x \neq y$, we must have either $x < y$ or $y < x$.
An equivalent formulation is the following:

For any x, y in W, either $x = y$ or $x < y$ or $y < x$.

This statement is known as the *Trichotomy law for* $<$. The word *law* is used simply for any true, general statement.

Sometimes the trichotomy law is expressed in a stronger form, as follows:

Trichotomy law for $<$: *For any* x, y *in* W, *either* $x = y$ *or* $x < y$ *or* $y < x$ -- *and no two of these three conditions can hold simultaneously*.

Another important fact about the ordering relation, less than ($<$), is the

Transitive law for $<$: *If* x, y, z *are any elements of* W *such that* $x < y$ *and* $y < z$, *then we must also have* $x < z$.
The truth of this can be seen from the meaning of $<$. For if $x < y$ and $y < z$, this means that in generating all whole numbers from 0 by successive applications of S, the number x comes before y and y comes before z. But then of course x is produced before z, which means $x < z$ as asserted in the transitive law.

There are still other laws which connect the successor operation, S, with the ordering relation, $<$. For example, *if* x *is any whole number then* $x < S(x)$. Again: *If* x, y *are any elements of* W *such that* $x < y$, *then also* $S(x) < s(y)$. *Con*versely, whenever $S(x) < S(y)$ we have also $x < y$.

Suppose that A and B are sets such that $n(A) < n(B)$. Then we say that A has fewer elements than B. Thus the relation $<$ between whole numbers provides a way of comparing the size of two given sets -- providing, of course, those sets are finite, so that the number of elements in each can be expressed by a whole number.

From the relation less than, $<$, and the relation of equality, $=$, we can define a new relation which is often quite useful in mathematics.

Definition of $\leq$. If x, y are any whole numbers we say that $x \leq y$ (read x is less-than-or-equal-to y), if either $x < y$ or $x = y$.

For example, $3 \leq 5$ and $3 \leq 3$, but not $3 \leq 2$.

The new relation also obeys a transitive law. Let us formulate this as a theorem.

Theorem. Using the transitive law for $<$ as a hypothesis, we can obtain as a conclusion the transitive law for $\leq$: If x, y, z are any elements of W such that $x \leq y$ and $y \leq z$, then also $x \leq z$.

To prove this, we begin by assuming that x, y, z are any whole numbers such that $x \leq y$ and $y \leq z$. By definition of $\leq$, this means that either $x < y$ or $x = y$, and also that either $y < z$ or $y = z$. Thus we get 4 possible cases for our assumptions:

Case 1. $x < y$ and $y < z$.
Case 2. $x < y$ and $y = z$.
Case 3. $x = y$ and $y < z$.
Case 4. $x = y$ and $y = z$.

We have to prove that in every one of these 4 cases we must have $x \leq z$. That is, we must show that in each of these 4 cases we have either $x < z$ or $x = z$. As a matter of fact, it turns out that in case 1, 2, and 3 we have $x < z$, while in case 4 we have $x = z$. The indicated conclusion for case 1 is immediate from the transitive law for $<$, and that for case 4 is immediate by the logic of $=$.

Actually, cases 2 and 3 can also be handled by the logic of $=$. Consider, for example, case 2, where our assumptions are that $x < y$ and $y = z$. This means that x is produced before y when the whole numbers are generated from 0 by successive applications of S, and that y is the same number as z. But then of course x is produced before z, which means $x < z$ as claimed in case 2. (See item 1, § 9.)

Other laws which hold for $\leq$ are as follows:

(i) For any x,y in W we have $x \leq y$ or $y \leq x$.

(ii) If x,y are any elements of W for which we have both $x \leq y$ and $y \leq x$, then we must have $x = y$.

The relation $\leq$ has a natural significance in connection with the counting process. To explain this, we must first define what it means for one set A to be a subset of another set B.

Definition. A set A is said to be a subset of another set, B, in case every element of A is also an element of B. To indicate that A is a subset of B we use the notation $A \subseteq B$.

Example: We have {London, Paris} $\leq$ {Paris, Milan, London} and

{2, 3, Moscow} $\subseteq$ {2, 3, Moscow}, but *not*

$$\{2, \text{London}\} \subseteq \{3, \text{London}, \text{Paris}\}.$$

Now a fundamental connection between the relation $\leq$ and the counting process can be described as follows:

Proposition. *Whenever A and B are (finite) sets such that $A \subseteq B$, then we must have $n(A) \leq n(B)$.*

To see this, suppose that $A \subseteq B$. Imagine that we count the elements of B, *starting* by first counting all of the elements which are in A before proceeding to any elements of B which are not in A. In this process, the number attached to the last element of A to be counted will be, of course, the number $n(A)$, since every element of A is in B by our supposition $A \subseteq B$. If there remain elements in B not yet counted, this means that $n(A)$ comes before $n(B)$ in the counting process so we have $n(A) < n(B)$. If, on the other hand, there are no elements of B left after those of A have been counted, then the last element of A to be counted will also be the last of B and so we have $n(a) = n(B)$. In either of these cases we have $n(A) \leq n(B)$, by the definition of $\leq$. (See § 9, Exercise 6.)

§ 8 (Track A)

In closing this chapter, let us draw attention to the many general statements (or laws) which we have cited. Why is the mathematician so interested in general statements? Actually, there are several reasons.

First of all, the discovery and communication of such laws provides mathematicians with a kind of esthetic satisfaction. Most people are not generally aware of this aspect of mathematical work. However, to discover that something is true of all elements of an infinite set, such as W, gives us insight into a kind of regularity among the elements of the set which is not unrelated to the element of form in a work of art.

General statements are also obviously economical ways of codifying many separate, independent, particular facts which would otherwise have to be registered separately.

A very important use for general statements is in connection with the deductive method for organizing our knowledge, which often leads us to discover new, particular facts about the domain under investigation. For instance, starting with a few general laws which we may know, or which we may assume as axioms, and using a few particular facts which may be known to us, or may be given as hypotheses, we can often combine these by the laws of logic into a proof whose conclusion may be some new fact, previously unknown, or not perceived as related to the given facts.

Finally, in the domain of whole numbers we shall see that the general laws are decisive in providing justification for the algorithms by means of which we all learn to carry out computations in elementary arithmetic.

§ 9 (Track B)

1. *The logic of* = . The reader who fills in the missing steps for the proof of the transitive law for the relation $\leq$ may notice that a transitive law for = is involved. He may well wonder why this was not explicitly listed as a hypothesis. This was not done since those laws dealing only with the relation = are considered parts of logic, an elementary working knowledge of which we are assuming the reader possesses.

2. *Two special subsets*. Let B be *any* set. Then we have both $B \subseteq B$ and $\emptyset \subseteq B$; in other words, every set is a subset of itself, and the empty set is a subset of every set. The first of these is easy to see from the definition of $\subseteq$, since of course every element of B is an element of B. To see the second statement, $\emptyset \subseteq B$, suppose that some set A is *not* a subset of B, i.e., that some element of A is not an element of B. Thus *if* we had *not* $\emptyset \subseteq B$ this would mean that some element of $\emptyset$ is not in B -- an impossibility, since $\emptyset$ has no elements whatever. Since it is impossible to have *not* $\emptyset \subseteq B$, we must have $\emptyset \subseteq B$ as claimed.

3. *Exercise*. a) Let $A = \{1, 2, 3\}$. Keep in mind item 2, above, and find *all* possible subsets of A. How many are there?

b) If B is any set such that $n(B) = 4$, how many subsets does B have altogether?

c) If C is a set with $n(C) = x$, where x is some whole number, generalize (b) above by finding how many subsets C has. (Express your answer by a formula involving the letter "x".)

4. *Other ordering relations on W*. Besides the relations $<$ and $\leq$ there are two other closely related relations which are used by mathematicians: $>$ (greater-than), and $\geq$ (greater-than-or-equal-to). The definitions are as follows:

Definition. *For any whole numbers x,y we define* $x > y$ *to mean that* $y < x$, *and we define*

$$x \geq y \text{ to mean that } y \leq x.$$

The reader can easily see that laws for $>$ and $\geq$ hold which are similar to those holding for $<$ and $\leq$ respectively.

5. *Exercise*. Prove the following law connecting the relations $>$ and $\leq$:

For any x,y in W we have $x > y$ *if, and only if, it is not the case that* $y \leq x$.

6. *Exercise*. Find a general law involving the operation S and the relation $<$, and illustrate it with a couple of particular examples.

7. *Exercise*. It was stated in Section 7 that if A and B are any sets such that $A \subseteq B$ then $n(A) \leq n(B)$. Consider the converse statement: If A and B are any sets such that $n(A) \leq n(B)$, then $A \subseteq B$. Is this general statement true or false? Explain why.

8. *Exercise*. In §7 we gave an argument to support the proposition that whenever we have $A \subseteq B$ we also have $n(A) \leq n(B)$. Give a similar argument to support the following proposition: If A,B are any sets such that $n(A) \leq n(B)$, then there is a subset C of B such that $n(C) = n(A)$. How would you find such a set C?

9. Exercise. Justifying your answer, indicate whether or not the set relation $\subseteq$ satisfies:

i) a transitive law

ii) a trichotomy law.

10. Classroom discussion. Discuss exercise 9 above. The discussion may touch on examples of other relations obeying one or both of these laws, on some general notions about order relations, or about still more general relations.

11. Exercise. Below there is stated the hypothesis and conclusion of a theorem, together with a sketch of a proof. Supply the missing parts of steps (3), (4) and (6), and the missing parts of the justifications for steps (2), (4), (5) and (7). Justifications may include references to theorems or propositions given in the text.

Theorem. Hypotheses: (i) A and B are sets with $x = n(A)$ and $y = n(B)$,

(ii) c is not an element of A, and d is not an element of B,

(iii) $A \cup \{c\} = B \cup \{d\}$.

Conclusion: $x = y$.

Proof: (1) $x = n(A)$, $y = n(B)$; hypothesis (i)

(2) c is not an element of A, and d is not an element of B; ______________

(3) ______________ $= B \cup \{d\}$; hypothesis (iii)

(4) $n(A \cup \{c\}) = S(x)$ and $n(B \cup \{d\}) =$ ________; from steps (1) and (2), by ________________.

(5) $n(A \cup \{c\}) = n(B \cup \{d\})$; by logic of = from step ____________.

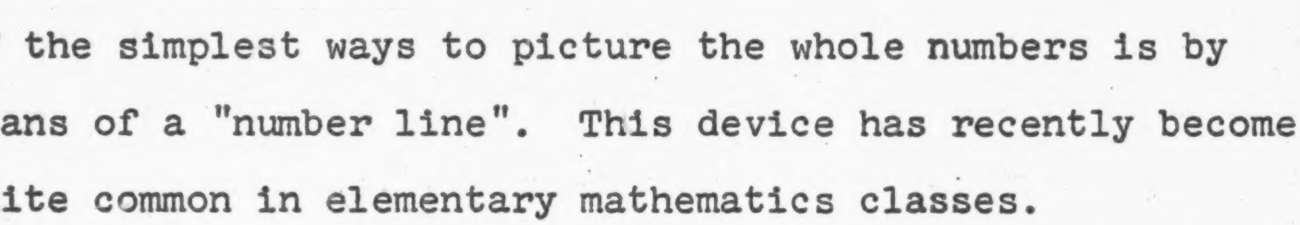
(6) _______________ $= S(y)$; by logic of = from steps (4) and (5).

(7) $x = y$; from step (6) since the operation S is ________, according to a proposition in § 1.

§ 10 (Track A)

Mathematicians like to illustrate the ideas of their theories with geometric diagrams of one sort or another, and the theory of the number systems is no exception. One of the simplest ways to picture the whole numbers is by means of a "number line". This device has recently become quite common in elementary mathematics classes.

It will be recalled that the geometric concept of a straight line is such that a line is considered to extend indefinitely in opposite directions, so that it has no ends. Of course when we draw a picture of a line we only represent a part of it, since in practice we must place our pencil down on a first point of the drawing and remove it from a last point. Examples:



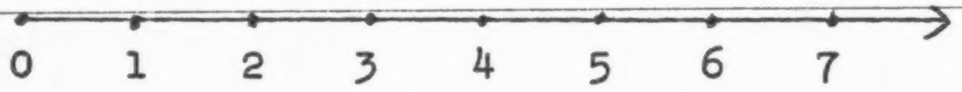
Technically, such a part of a straight line is called a *line segment*.

Actually, for purposes of picturing the whole numbers we do not require a whole line, but only a half line (Also called a *ray*). Customarily we draw this horizontally, proceeding to the right of our starting point, thus:

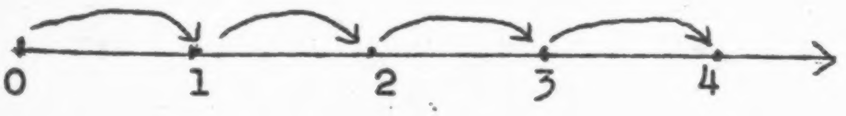
We place the arrow head at the right end of our drawing to indicate that we wish to consider the unending half line which continues indefinitely to the right. Sometimes, however, we shall find it convenient to draw a number line pointing in some other direction, such as upward.

Now we indicate certain points on our half line by means of small circles, large dots, or little cross marks. The left end-point of the ray is one of these marked points, and we label it with the numeral 0. We then choose some other point of the ray to the right of the end point, and label it 1. We often refer to the points bearing these labels as "the point 0" and "the point 1". The distance between these two points is called *the unit distance* of the number line.

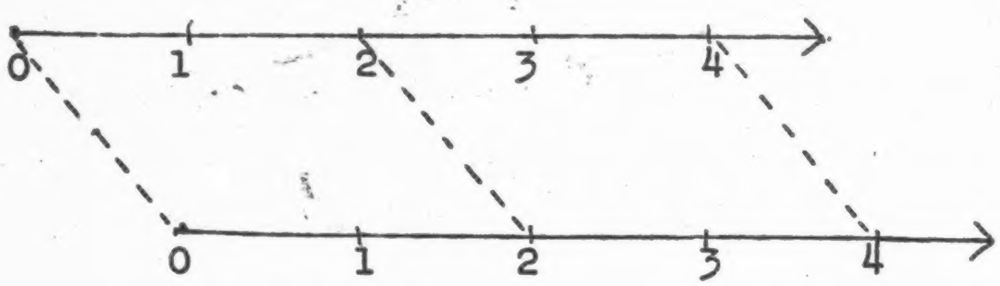
Now, starting at the point 1, we lay off to the right of it the unit distance, arriving at a new point which we label 2. Then the same distance is laid off to the right of 2 to reach a new point, labeled 3. Of course on a given picture we can generally fit only a few points. However, since the ray proceeds indefinitely to the right, we can imagine that every whole number is assigned as a label to some point of the ray. The ray, together with the special labeled points, is called the whole number line. A picture would look like this:



With the aid of such a whole number line we can picture the operation S and the relation < in various ways. For instance one way to picture the operation S is to draw little curved arrows above the line, starting from each numbered point x and finishing at S(x). This would look as follows:



These curved arrows suggest a motion of the line, in which the line moves one unit distance to the right. We can illustrate this motion by drawing the number line in its initial position, and then under it we draw the line in the position it would occupy at the end of this motion, as follows:

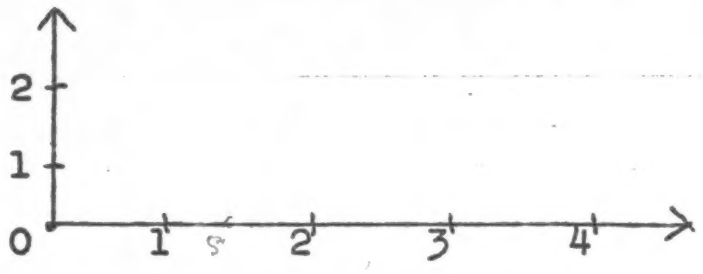


From such a double diagram we can read off the successor of a number on the lower line simply by looking at the number right above it on the upper line.

As for picturing the ordering relation, <, this is most simply done by noticing that whenever x and y are numbers such that x < y, then the point labeled x is to the left of the point labeled y.

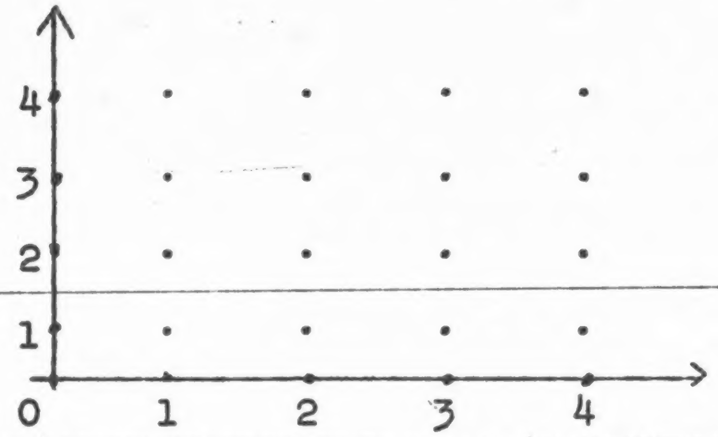
We would like to emphasize that there is no unique way to picture a mathematical concept. For example, both the successor operation S and the ordering relation < can be pictured in ways quite different from those indicated above by means of the notion of graphs, which we now proceed to explain.

As the framework for a graph we construct two number lines, starting from the same point, at right angles to each other. It is customary to draw one of these horizontally to the right and the other vertically upward, as follows:



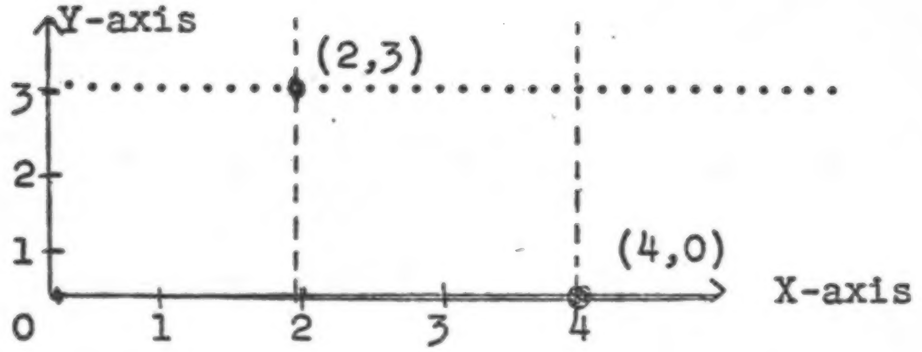
These two number lines are called the coordinate axes of the graph.

Now let us imagine all of the vertical lines which pass through the labeled points of the horizontal axis, and all of the horizontal lines which pass through the labeled points of the vertical axis. The points where these vertical lines intersect these horizontal lines are called the lattice points of the graph. They are pictured as follows:



Notice that the labeled points on the axes are themselves among the lattice points.

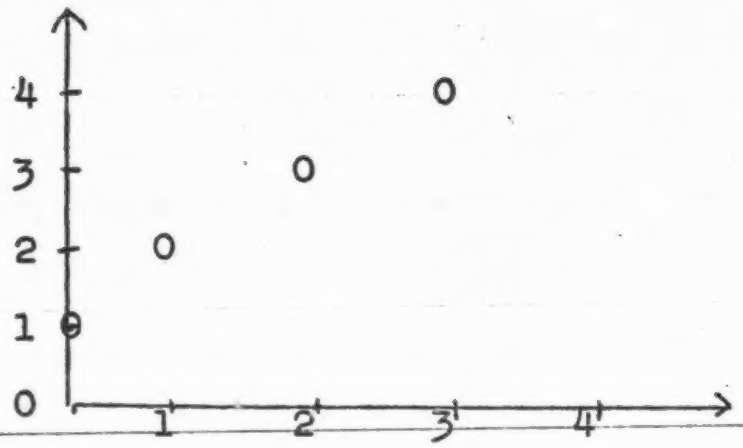
Now with any lattice point there is associated an ordered pair of whole numbers, (x,y). We get the first number, x, by following down the vertical line through the lattice point and seeing which labeled point on the horizontal axis lies at its foot. (The horizontal axis is sometimes called the X-axis.) And we get the second number, y, by looking across the horizontal line through the lattice point and seeing which labeled point on the vertical axis lies at its left end. (The vertical axis is sometimes called the Y-axis.) The numbers x,y associated in this way with a given lattice point are called the coordinates of the point, and are often written near the point on a picture as a kind of label for that point. We often speak of the point which is labeled with the number pair (x,y) as "the point (x,y)". Pictured below are the points (4,0) and (2,3).



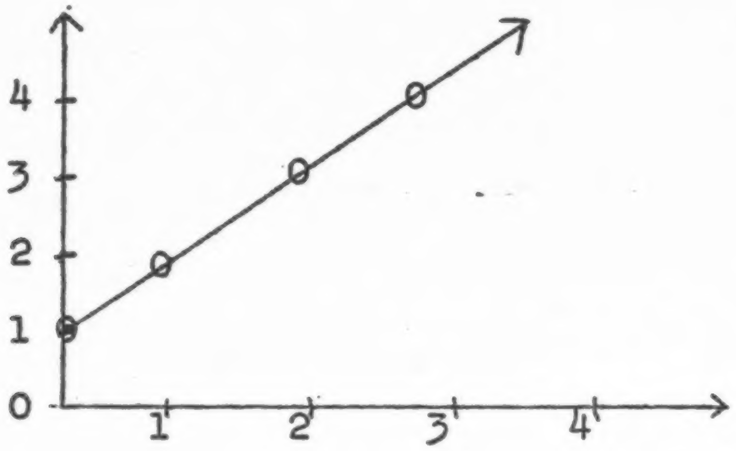
It is easy to see that not only does every lattice point have a pair of coordinates, (x,y), but that conversely, given any pair (x,y) of whole numbers, we can always find a lattice point having those numbers as its coordinates.

Now let us consider the equation $y = S(x)$. If we consider any ordered pair of whole numbers, such as (2,3) or (2,4), we can substitute the numbers of the pair for the letters of the equation, always following the rule that the first number of the ordered pair is substituted for the letter "x", and the second number is substituted for the letter "y". The result of such a substitution is a particular statement about the successor operation which may be true or false. For example, substituting (2,3) we get the true statement $3 = S(2)$, while substituting (2,4) gives the false statement $4 = S(2)$. An ordered pair of whole numbers is said to satisfy the equation if the substitution results in a true statement.

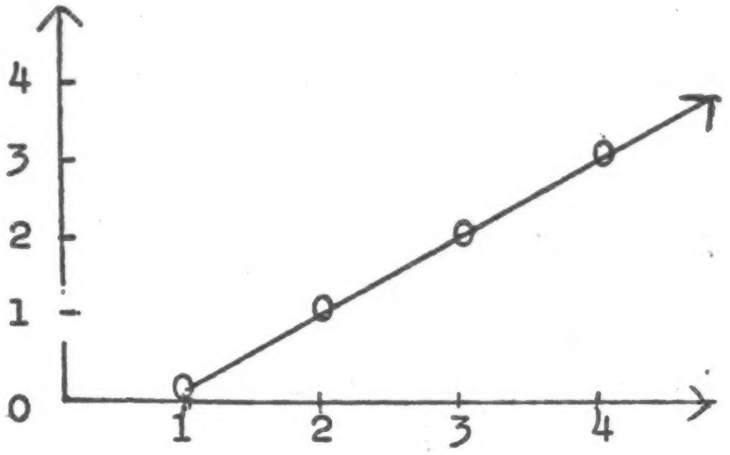
Now the graph of the equation $y = S(x)$ is simply the totality of lattice points whose coordinates satisfy the equation. Thue the graph consists of all the infinitely many lattice points (0,1), (1,2), (2,3), (3,4), ... Of course we only picture a few of these.



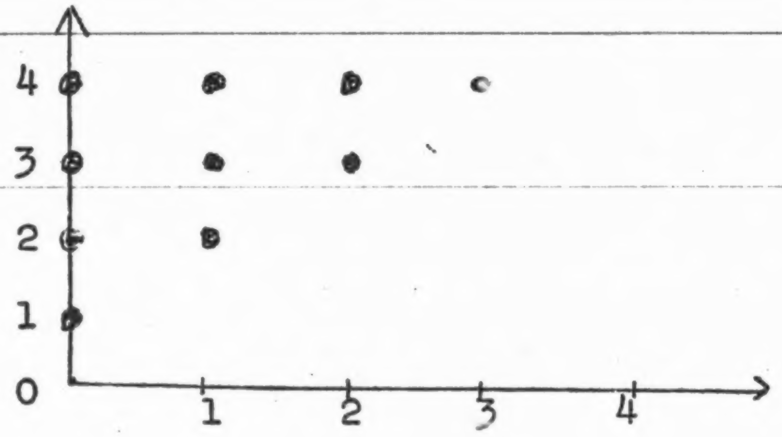
It is evident that all of these points lie on one straight line, and we often join them by a pictured line.



In the same way we can picture a graph for the equation $x = S(y)$. Remembering our rule of substituting the first number of an ordered pair for "x", the second for "y", we see that (3,2) will satisfy this equation, but (2,3) will not. Here is a picture of the graph of $x = S(y)$:



We can also draw graphs for an inequation, such as $x < y$. We see that (2,4) satisfies this inequation because the substitution of "2" for "x" and "4" for "y" yields a true statement, but (4,3) and (2,2) do not satisfy it. Here is a picture of a portion of the graph of $x < y$.

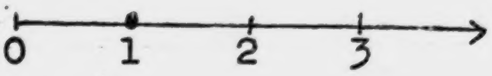


We see that the graph of $x < y$ consists of the graph of $y = S(x)$ together with all lattice points lying above the latter. This relationship between the graphs gives us a visual image corresponding to the relationship between the operation S and the relation < .

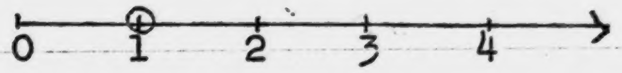
§ 11 (Track B)

1. Graphs on a number line. In §10 we introduced graphs of equations such as $y = S(x)$, and inequations such as $x < y$, involving two letters (or variables, as they are called). Even simpler equations and inequations, involving only a single variable, also can be represented pictorially. Whereas the graph of an equation with two variables is a set of points in a plane, the graph of an equation involving one variable is a set of points on a single number line.

For example, consider the equation $2 = S(x)$. There is only one whole number which satisfies this equation and that is 1. Hence the graph of $2 = S(x)$ is represented by a single point.



We generally picture such a graph by making a heavy dot at points of the graph, or by circling these points.



2. __Exercise.__ (a) On a whole number line graph the equation $x = S(2)$. (This means, draw a picture of the graph of the equation.)

(b) Graph the inequation $x < 4$.

(c) Graph the inequality $x \neq S(2)$. (Circle the points of the graph, using two "circles".)

(d) Explain the connection between the Trichotomy law for the relation $<$ and the fact that the graph in (c) falls naturally into two parts.

3. __Exercise.__ Using coordinate axes, graph the following equation which uses the identity operation: $y = I(x)$. (See item 1, §3.)

4. __Exercise.__ Using a single set of coordinate axes, graph the equations (a) $y = S(x)$

(b) $y = S(S(x))$

(c) $y = S(S(S(x)))$

What is the relation of these graphs to the graph of $x < y$?

2

5. Classroom discussion. The relation "to the left of" on the number line is a pictorial representation of the relation $<$ on the set W of all whole numbers. However, if we look at the lattice points determined by a pair of coordinate axes, they do not appear to be "lined up" in any natural way. But let us try to order, or "line up", these points.

At the board, draw a picture of coordinate axes and lattice points, and have several students try to describe one or more orderings of the lattice points. This means describing a rule which tells when a point (x,y) comes "before" some other point (u,v). To satisfy the trichotomy law, we must be sure that for any two distinct lattice points, one of them comes before the other. In other words, if we use the notation $(x,y) \ll (u,v)$ to indicate that the point (x,y) comes before the point (u,v), we must be sure that whenever $(x,y) \neq (u,v)$ we have either $(x,y) \ll (u,v)$ or $(u,v) \ll (x,y)$, but not both. The ordering may be described with the help of pointing, but the rule defining $\ll$ must cover all the infinitely many lattice points, not just the ones pictured on the board.

Discuss in what ways the ordering $\ll$ of lattice points resembles the ordering $<$ of points on the whole number line, e.g., the transitive law. Discuss in what ways the two ordering relations may differ, e.g., is there more than one lattice point which lacks an immediate predecessor with respect to the ordering $\ll$? (If so, can we define some other orderings of the lattice points in which there is exactly one such point, or in which there are none?)

After an ordering << has been defined "pictorially", i.e., with the help of pointing, an attempt can be made to give a mathematical description in terms of the number pairs (x,y) which are used as labels for the lattice points.

Among the natural orderings likely to be devised are the lexicographical orderings of the number pairs induced by the relations < and > on W. (See Ex.7, Chap1, §5.) If these have not been found by the students, the instructor should define the lexicographical ordering << induced by < and then ask for a pictorial account of the relation << in terms of vertical lines on the coordinate plane. Afterward, raise the question of horizontal lines.

6. Exercise. In each of (a), (b), (c) below there are given the coordinates of three lattice points (x,y) which lie on one straight line. In each case plot the points and connect them with a straight line. Then try to find an equation involving the letters "x" and "y" which is satisfied by the coordinates of all three of the given points. Finally, find a fourth point whose coordinates satisfy the equation you found, and determine whether or not this point lies on the straight line passing through the other three points.

(a) (0,0), (1,2), (3,6)
(b) (3,1), (0,0), (6,2)
(c) (0,4), (2,2), (3,1)

7. **Exercise**. In each of (a), (b), (c) below there are given the coordinates of two lattice points (which of course determine a straight line). Find the coordinates of two more lattice points on each of these lines.

(a) (0,0), (3,2)
(b) (0,0), (2,3)
(c) (0,1), (4,5)

8. **Exercise**. Consider a geometric plane (without any coordinate axes specified). If we mark off any two points in this plane they can be connected by a straight line segment (i.e., a piece of a straight line whose ends are at the given points). If we have three points marked in the plane which do not all lie on one straight line, and if we wish to connect each marked point to both of the others, we need three segments. If we have four marked points in the plane, no three of them on a single straight line, then six segments are needed to connect each marked point to all of the others. (Make a drawing to show this.)

Collecting the above information in a table, we have:

No. of points	No. of connecting segments
2	1
3	3
4	6

Now suppose that n is any whole number. Try to find a formula, which may involve the letter "n" and symbols for any of the arithmetical operations, which gives the number of segments needed to connect each of n marked points to all of the others (assuming no three of the marked points lie in one straight line). Check your formula with a drawing for the case $n = 5$.

9. Exercise. Read the dialogue in item 2 of §12 below. Then write a dialogue to illustrate the type of game described in item 3 of §12.

10. Exercise. Write out an explanation of how, when we are given a pair of coordinate axes and two whole numbers x, y, we can find the point labeled (x,y) on a graph.

§12 (Track C)

1. Importance of Introducing the Whole Number Line. The whole number line is very useful at the primary level for illustrating the whole numbers as an ordered set arranged in increasing order from left to right, if the line is oriented in the normal horizontal position, or from bottom to top if given the normal vertical orientation. It should be introduced as soon as the names for the whole numbers have been learned since it can represent pictorially concepts which are much more difficult to introduce verbally. For example, on the number line it is immediately evident that for any two distinct whole numbers, one of them must be less than the other (Trichotomy law). Similarly the transitive property of the relation $<$ can be easily understood visually.

2. Games using the Whole Number line. Example. A number line is drawn on the chalk-board.

Teacher: I'm thinking of a number between 0 and 10. I'll call it x. Can you guess which number x is?

Child: Is it 9?

T: No, x is less than 9. (Writes $x < 9$ on board.)

C: Is it 1?

~~T: No, x is more than 1. (Writes x > 1.)~~

C: Is it 10?

T: No, because x is less than 9. (Points to "$x < 9$" on board.) If the child still doesn't understand, have him go to the board and show on the number line which numbers are less than 10.

C: Is it 5?

T: No, x is less than 5. (Writes $x < 5$.)

C: Is it 2?

T: No, x is more than 2. (Writes $x > 2$.)

C: Is it 3?

T: Yes, and since you guessed it, it's your turn to be the teacher and we'll try to guess what number you're thinking of. (Since young children can forget with ease, it would help to have him write the number on a slip of paper.)

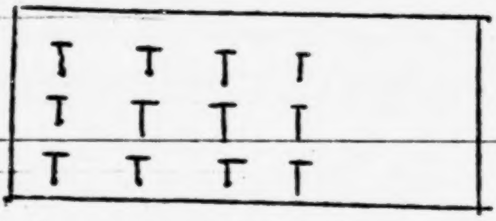
After the game the teacher could point to the written statements $x < 9$ and $x < 5$. The former can be crossed out and the teacher can ask why the information in that statement is no longer needed, once $x < 5$ has been written down.

One can vary the game in many ways. An objective to work towards is having the children not try to guess the number, but instead to ask questions that would narrow in on the hidden number. One way to do this would involve only answering yes-or-no type questions, and perhaps imposing some limit on the number of questions that could be asked. As children become more sophisticated they might be asked to try to determine if there is some minimum number of questions that would guarantee them getting the hidden number.

3. Guessing games in the plane. Guessing games of the above sort are also useful with reference to a pair of coordinate axes. Here of course the aim is to guess the coordinates of a point. Hence answers to questions would involve the relation (larger or smaller) of each coordinate of the guessed print to the corresponding coordinate of the hidden point.

4. Drawing graphs in elementary school. Plotting points on a coordinate axis system requires practice for children. One way to make it more interesting would be to have them plot points and connect them to make some kind of closed figure such as a triangle, square or other figure. As they achieve proficiency at this, they can be asked to make their own simple figures on graph paper, and then to make up a table of coordinates of the vertices of their figure. Each student can then exchange these tables of coordinates with his neighbor, who would then try to reconstruct the original figures.

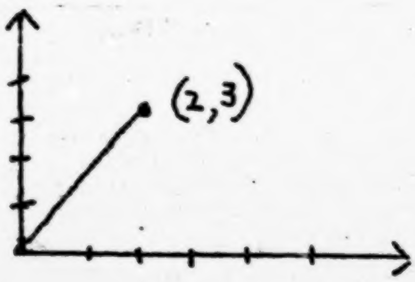
In this connection we should call attention to "geoboards". These are commercially available, but can also be made easily by youngsters -- and even, in some cases, by teachers! A geo-board is a board into which upright nails have been hammered so as to form a rectangular array.



These boards are used with rubber bands. Such a band can be stretched between two nails to form a line segment, or it can be stretched around several nails to form a triangle or other polygon -- convex or other.

At a later stage these can be used to illustrate many mathematical ideas, for instance those connected with area. However, they can be used very early for practice in describing an array of points by means of coordinates.

5. Using a coordinate system drawn on the chalkboard, draw a line segment connecting the point (0,0) to any other lattice point (see below). Ask the students to find the coordinates of another point on this line. Then see if they can find something in common about the coordinate-pairs of all the labeled points on the line. Have them guess what the coordinates at the next lattice point would be if the segment is extended. Similarly, start with the segment joining (1,2) with (3,4) and ask for other points on the line, and for something which all coordinate-pairs on that line have in common.



Chapter 3. Addition.

§1 (Track A)

In the preceding chapter we have examined two aspects of the whole number system which are basic to the use of these numbers in counting: the successor operation, S, and the ordering relations $<$ and $\leq$. In the present chapter we shall study the two-place operation of addition, $+$, on the set W of all whole numbers.

This operation is usually introduced in the elementary schools in terms of the counting process. In order to add two given whole numbers x, y a child is instructed to begin by taking two sets A and B, the number of elements in these sets being x and y respectively. (Very often the elements of A and of B are the child's own fingers.) He then combines the elements of these two sets into one large set -- mathematically speaking, he forms the union $A \cup B$ of the two chosen sets -- and counts the number of its elements. The resulting number, $n(A \cup B)$, is defined to be the sum $x + y$ of the two given numbers.

Basically, this is a perfectly satisfactory way of obtaining a mathematical definition of addition, providing we take care of two details. The first of these is that we must make sure that the two chosen sets, A and B, are *disjoint*, i.e., that they have no common elements. Ordinarily this point is not mentioned explicitly in the elementary school, or at any rate is not stressed, since the teacher can rely on the particular sets A and B turning out to be disjoint just from the way in which they are chosen -- e.g., A may consist of fingers from the left hand, B of fingers from the right hand, and so of course the two sets will have no common elements. It is intuitively clear, however, that if the sets A and B have several common elements then the combined set, $A \cup B$, will have fewer than $x + y$ elements in it.

How does the mathematician symbolize the fact that two sets, A and B, have no common elements? In the first place, given any sets A and B whatever, there is a notation for the set of all their common elements: $A \cap B$. This set, consisting of all those objects which are in both A and B, is called the intersection of A and B. Now, to say that two given sets have no common elements is simply to say that their intersection is empty. Recalling our symbol $\emptyset$ for the empty set, we see that to express the fact that A and B are disjoint we can write: $A \cap B = \emptyset$.

We have observed above that if $n(A) = x$ and $n(B) = y$, then the condition $A \cap B = \emptyset$ must be satisfied if we are to have $n(A \cup B) = x + y$. But there is a second point we must consider in connection with the elementary-school way of introducing addition, before we can base a precise mathematical definition upon it. This second point is somewhat more subtle, so let us examine it in detail.

Starting with given numbers x,y of the set W, a child chooses disjoint sets A and B so that the number of elements in A is x (i.e., so that $n(A) = x$), and so that $n(B) = y$. The child then forms the union, $A \cup B$, counts the number of elements $n(A \cup B)$, and is told that this number is $x + y$. Of course another child in the class, when the same numbers x and y are given, will generally not choose the same sets A and B as the first child -- especially if each child is taught to make up the chosen sets using his own fingers as elements! So the second child chooses a different pair of disjoint sets, say C and D, with $n(C) = x$ and $n(D) = y$; he then forms the union of his sets, $C \cup D$, counts its elements getting $n(C \cup D)$, and is told that this number is the required sum, $x + y$. But how do we know that the number of elements in $C \cup D$ is the same as the number in the set $A \cup B$ which the first child is counting? True, A has the same number of elements as C -- this number being x . And B has the same number as D -- this being the number y . So we expect that $A \cup B$ will turn out to have the same number of elements as $C \cup D$. But why do we expect this outcome? And is an expectation the same thing as mathematically certain knowledge? One thing is clear -- if we ever had a case where $n(A \cup B)$ turned out to be different from $n(C \cup D)$, i.e., where $n(A \cup B) \neq n(C \cup D)$, then it wouldn't make sense to call both of these numbers $x + y$. Thus, from the mathematical viewpoint this method of defining addition depends upon a prior knowledge that in fact we will have $n(A \cup B) = n(C \cup D)$ in every case.

Let us formulate the desired result as a theorem.

Theorem. Let x and y be any whole numbers. Let A, B, C, D be any sets such that: (i) $n(A) = x$ and $n(C) = x$, (ii) $n(B) = y$ and $n(D) = y$, and (iii) $A \cap B = \emptyset$ and $C \cap D = \emptyset$. Then we must have $n(A \cup B) = n(C \cup D)$.

In elementary school we come to believe this result on the basis of experience. That is, we try it out using several choices for x, y, A, B, C, and D, and we find that in each case we get $n(A \cup B) = n(C \cup D)$ -- or at any rate, if we don't, the teacher tells us we must have made a mistake!

However, a mathematician is never satisfied to establish a statement about all numbers x and y, or about all sets A, B, C, and D, just by trying a few cases. He wants to know whether it is possible to prove that the desired result will be obtained in every case.

Of course every proof must start with some facts, or principles, which are used in the proof. A proof of the above theorem can be given, based upon the principle of mathematical induction. (We have mentioned this principle in Chapter 2, but without formulating it explicitly; we shall defer a formulation until the end of the present chapter, §9.) Once the theorem is established, we are justified in formulating the following definition of addition.

First Definition of Addition on W. We define $+$ to be the two-place operation on W such that, given any numbers x,y of W, the number $x + y$ (called the sum of x and y) is obtained by first choosing any disjoint sets A, B such that $n(A) = x$ and $n(B) = y$, and then letting $x + y$ be the number $n(A \cup B)$.

This definition is justified by the preceding theorem.

The reader may be surprised to see the word "First" in the title of our definition above. Can we have more than one definition for a given concept? Do we need more than one?

The answer is that when a mathematician is developing a particular mathematical theory he needs only one definition for any particular concept he wishes to introduce. Similarly, he needs only one proof for any particular theorem he wishes to establish. However, in general he has a choice of more than one definition, or more than one proof, which he can use at each stage. It is important for a mathematics teacher to realize that there is almost never a unique way to solve a given mathematical problem or attain a given mathematical goal. One of the most widespread faults in elementary mathematics teaching is the insistence that all students do a given type of problem in a rigidly prescribed manner. This stifles the processes of exploration and discovery, and the creative activity of developing original ways of doing things, which provides much of the excitement of mathematics. Of course, to encourage such creativity a teacher must know how to recognize a mathematically satisfactory answer to a given problem even when it is different from the solution worked out in the teacher's edition of the textbook. To help develop such an ability we shall often exhibit more than one possible way of defining a given concept, or more than one way of proving a given result, in these notes.

Our first definition of addition was based on the use of the whole numbers in counting the elements of certain sets. Let us give another definition now in terms of the successor operation. The reader may recall that multiplication is often introduced as "repeated addition" in elementary school; we will examine this in the next chapter. The point we want to make at this time is that addition also can be introduced as the repeated use of some other operation -- namely, of the successor operation, S .

Second Definition of Addition on W. We define + to be the two-place operation on W such that, given any numbers x, y of W, the number x + y (called the sum of x and y) is obtained by applying the operation S successively y number of times to the number x . In symbols:

$$x + y = S\underbrace{(\ldots.S(S(x))\ldots.)}_{y \text{ applications of } S.}$$

In case y = 0 we apply S no times, so that x + 0 = x.

For example, to compute $2 + 3$ by this definition, we would compute

$$\begin{aligned} S(S(S(2))) &= S(S(3)), \text{ since } S(2) = 3 \\ &= S(4), \text{ since } S(3) = 4 \\ &= 5 . \end{aligned}$$

To compute $2 + 3$ by the First Definition we would first choose sets A and B, say A = {George Washington, Abraham Lincoln} and B = {Paris, London, Bangkok}, such that $n(A) = 2$ and $n(B) = 3$ and $A \cap B = \emptyset$. We would then form $A \cup B$, namely,

{George Washington, Abraham Lincoln, Paris, London, Bangkok}, and finally we would count the number of elements in this set, $n(A \cup B)$, getting the same answer 5 as we got above when we used the Second Definition. The fact that our two definitions of $+$ will give the same answer for every pair of whole numbers x,y is something a mathematician would wish to prove, but we shall simply take it for granted in our present treatment of the subject.

Using either definition of $+$, we can establish a great many particular statements about this operation such as the fact that $2 + 3 = 5$, which was derived above. For example, we get such facts as $0 + 2 = 2$, $8 + 1 = 9$, $4 + 4 = 8$, etc. Often we collect together facts of this kind into a table of values of $+$, or an addition table. For example, a table giving all sums $x + y$ for $x,y = 0, 1, 2, 3$ would look like this:

+	0	1	2	3
0	0	1	2	3
1	1	2	3	4
2	2	3	4	5
3	3	④	5	6

To find $3 + 1$ in this table we locate 3 in the left margin, 1 at the top margin, and find the entry which is across from the 3 and under the 1, namely, the circled 4. So $3 + 1 = 4$.

In the elementary schools it is customary to memorize the facts contained in a table of this kind. For example, one often memorizes the value of all sums $x + y$ where x,y are any of the numbers $0, 1, \ldots, 9$. However, no one can write or memorize the values of $x + y$ for all whole numbers x,y because there are infinitely many of them. What we do in practice is to use the (memorized) table of sums of one-digit numbers, together with a certain procedure (or algorithm, as it is called). Both of these together enable us to compute the sum of any given two whole numbers. One of the principal thrusts of recent elementary mathematics curriculum revisions has been to get across to students an understanding of why these algorithms lead to a correct computation of sums in every case. We shall see that from the theoretical point of view such an understanding depends critically on the general statements which are true about the operation of addition. Let us, therefore, turn our attention to some of these.

§ 2 (Track B)

1. Class Discussion. According to the First Definition of addition on W, finding the sum of two whole numbers x and y first requires the choosing of two disjoint sets A and B such that $n(A) = x$ and $n(B) = y$. If x and y are small enough this choice is easy, for ones own fingers will do very well. However, what if x and y are very large? How can we construct sets large enough for addition of any whole numbers x and y?

2. Exercise. (a) Find two sets S and T such that $n(S \cup T) = n(S) + n(T)$.

(b) Find two sets W and V such that $n(W \cup V) < n(W) + n(V)$.

(c) Can one find sets W and V such that $n(W) + n(V) < n(W \cup V)$? Explain.

(d) Let $A = \{1, 2, 3\}$. Make a list of all sets B such that $A \cap B = B$. (There are 8 of them.) What do they all have in common?

(e) Letting $A = \{1, 2, 3\}$ as in part (d), how many sets B are there such that $A \cap B = A$?

(f) What property is common to all those sets B in part (e)?

3. Exercise. The theorem which justifies the First Definition of addition states that if A, B, C, D are any sets such that (i) $n(A) = n(C)$, (ii) $n(B) = n(D)$, (iii) $A \cap B = \emptyset$, and (iv) $C \cap D = \emptyset$, then we must have

(v) $n(A \cup B) = n(C \cup D)$. However, it is possible to have sets A, B, C, D which satisfy (v) even though (i), (ii), (iii), and (iv) are not all true. Find sets A, B, C, D such that $n(A \cup B) = n(C \cup D)$ and $A \cap B = \emptyset$, $C \cap D = \emptyset$, but $n(A) \neq n(C)$ (and also $n(A) \neq n(D)$). Is such an example possible if $n(B) = n(D)$? Why?

4. Exercise. Compute $0 + 4$, first using one of the definitions of addition given in §1 then using the other definition.

5. Project to be written up and subsequently discussed. Find a child around five or six years of age who is familiar with the names of at least the first five or six whole numbers. At this age most children have some ideas concerning very simple sums, such as "two plus two." Using as light a touch as possible, try to discover just what these ideas are. You will have to be understanding and skillful, for children this young are not always very verbal and are put off quite easily. Make an attempt at figuring out how your subject determines different sums. Do the two definitions of addition given in the text play any part in the child's conception of adding numbers? Write up your observations and comments and discuss them in class.

6. Class Discussion. In §1 two definitions are given for the operation of addition. Why do we not say that these two definitions define two distinct operations? Discuss a way in which one might see, intuitively, that the computation

of $x + y$ by the Second Definition will give the same result as by the First Definition, no matter what whole numbers x and y we have.

Suggestion: Starting with disjoint sets A and B having x number of elements and y number of elements respectively, we form $A \cup B$ (according to the First Definition) by combining the elements of A and B into one big set. Now suppose we carry out this combining process by carrying the elements of B one at a time over to A, thus enlarging A through a series of intermediate sets until $A \cup B$ is achieved. If we keep track of the number of elements in these intermediate sets, do we see any connection with our Second Definition of addition?

§3 (Track C)

1. Introducing Addition. One of the requisites for finding the sum of two whole numbers x and y, according to the First Definition of Addition, is finding two disjoint sets A and B with $n(A) = x$ and $n(B) = y$. The manner by which these two sets are chosen and counted is mathematically unimportant, but children should be brought to realize the importance of the fact that it doesn't matter which sets are selected -- a fact which was formulated as a theorem in §1.

In order to facilitate the discovery of this fact, however, when the subject of addition is first introduced, it is pedagogically desirable to have some standard objects which the children can easily use as elements of sets. For example, it would be useful to have a box of green counters marked from one to thirty, and another box of red ones marked similarly. By taking the sets A and B in the definition of addition to be sets of green and of red counters, respectively, and then forming a new set (the union), children can get their first practice in counting unions.

The teacher should realize, however, that there are different ways of counting the union of two sets of marked counters, and that by asking skillful questions the children can be led to discover various facts about this process and thus led to make the process more efficient.

For example, consider first the problem of obtaining x green counters, x being a number given by the teacher. The children can at first be asked to take x counters when these are placed blank side up (no numerals showing). Under these conditions the counting process must be carried through, of course. Afterward the same type of task can be set when the numerals marked upon the counters are showing. By asking (if necessary) whether the numerals can be of help, the students should be led (after experimentation) to realize that in order to secure x counters they need only take those counters marked 1, 2, ..., x -- without actually counting. The game can then be made more difficult by removing the counter marked with 1, say, or those marked with 1 and 2.

After the process of selecting a given number of colored markers has been improved, as above, attention should be turned to the process of counting a set consisting of x green counters and y red ones. If the green and red counters are mixed before the counting begins, then all the elements of the set must be counted. However, the teacher should then suggest that all the green counters in the set be counted first, before any of the red ones. The children will soon realize that there is no use in counting the green ones, for the answer will always be x -- the given number. Hence, when asked to count the union they may begin the counting process by touching the first red one and calling out $x + 1$, touching the second and calling out $x + 2$, and so on until the last red one is counted when the desired number, $x + y$, will be called out.

A further level of sophistication can be reached by taking x and y to be very unequal numbers, and asking the children to count the union in two ways -- first counting all the green counters before the reds, and then in the opposite order. They will presently be led to observe that it is easier if one takes the larger number first, since the counting process (as indicated in the previous paragraph) need only involve touching the elements of the second set.

In proper sequence the teacher can thus guide the children to successive levels of sophistication by using a series of directed questions, asking such questions as "Which set should we count first?" or "Does it matter which set we begin with?" or "What about the numerals painted on the counters; can we use them to help us count the sets A and B ?" In other words possible (and sometimes impossible) points of departure from a previously used procedure are called into focus for discussion and implementation (or rejection).

§4 (Track A)

The commutative law for + : For any x,y in W we have $x + y = y + x$.

This is so familiar and intuitively clear to most of us that some students have difficulty in seeing that the statement really expresses anything at all! Perhaps the way to begin is to notice that there are some two-place operations which do not obey a commutative law. Subtraction and exponentiation are examples -- while we will study these in later lectures, the reader has encountered them in elementary mathematics courses and knows that there are numbers x,y such that $x - y \neq y - x$, or such that $x^y \neq y^x$. Thus, the fact that addition and multiplication obey commutative laws sets these operations apart from others like subtraction and exponentiation.

The evident truth of the commutative law for + is seen clearly from our first definition of addition. For in order to compare $x + y$ and $y + x$, where x and y are any given whole numbers, that definition requires us first to choose disjoint sets, say C and E, such that $n(C) = x$

and $n(E) = y$. Then by the first definition of $+$ we have $x + y = n(C \cup E)$ and $y + x = (E \cup C)$. But $C \cup E$ is the same set as $E \cup C$, for in either case we form the set by combining the elements of C and of E into one big set. Since $C \cup E = E \cup C$, of course we get $n(C \cup E) = n(E \cup C)$ by the logic of $=$, that is, $x + y = y + x$.

It is worth noting, however, that the truth of the commutative law for $+$ is not equally self-evident if we use our second definition of addition. Being given two whole numbers x and y, in order to show that $x + y = y + x$ according to that definition, we would have to show that applying S successively y number of times to x, brings us to the same result as applying S successively x number of times to y. While we could check that this indeed turns out to be the case in individual instances by choosing a few special values of x and y, some further argument would be needed to prove the result in full generality. (Such an argument can be provided by using the principle of mathematical induction, §9.)

The associative law for $+$. For all x, y, z in W we have $(x + y) + z = x + (y + z)$.

Students often have difficulty in seeing the meaning of this law because they have been taught to add a column of three numbers in elementary school. Actually, the possibilit of adding together a column of more than two numbers, and getting the same result no matter whether we add from top to bottom or vice-versa, rests upon the associative law. It is important to realize that addition as originally defined is a two-place operation -- using either one of our definitions of +. Hence at first we can only add two numbers at a time. It follows that if someone gives us three numbers in a certain order, say x,y,z, the only way we can add them all is to apply the two-place operation + twice. But when the numbers x,y,z are given in this specific order there are two different ways of applying addition twice: One way is to form the sum $x + y$ and then add this number to z getting $(x + y) + z$; and the other way is to first form the sum $y + z$, and then add x to this number getting $x + (y + z)$. Will the numbers resulting from these two different ways of applying + twice, namely, the numbers $(x + y) + z$ and $x + (y + z)$, turn out to be the same -- no matter which numbers x,y,z we start with? The associative law gives an affirmative answer to this question.

The truth of the associative law for + can best be seen from our first definition of +, though it is not as easy to see as in the case of the commutative law. Being given any three whole numbers x,y,z, we first choose three sets A, B, C such that $n(A) = x$, $n(B) = y$, and $n(C) = z$ -- making sure that no two of the sets A, B, C have an element in common.

Under these conditions we will have, by our definition of +, that $x + y = n(A \cup B)$. Then, because of the way in which A, B, and C were chosen, and because $A \cup B$ is obtained simply by combining the elements of A and B into one big set, we see that the two sets $A \cup B$ and C will have no element in common. Hence we can apply our definition of + a second time and conclude that $(x + y) + z = n((A \cup B) \cup C)$. Since the set $(A \cup B) \cup C$ is formed by combining the elements of $A \cup B$ and of C into one big set, we see that in fact $(A \cup B) \cup C$ is obtained by combining all of the elements of the three sets A, B, and C into one big set.

By entirely similar arguments we can see first that $y + z = n(B \cup C)$, and then that $x + (y + z) = n(A \cup (B \cup C))$. Furthermore, $A \cup (B \cup C)$ turns out also to be the set obtained by combining all of the elements of the three sets A, B, and C into one big set. In other words, we find out that $(A \cup B) \cup C$ is the same set as $(A \cup (B \cup C)$. Obviously, then, $n((A \cup B) \cup C)$ is the same number as $n(A \cup (B \cup C))$, i.e., $x + (y + z) = (x + y) + z$.

The truth of the associative law can also be established using the second definition of +, but as in the case of the commutative law a very different kind of argument is needed. (For a proof of the associative law from certain axioms, see §7 below.)

The following two laws involve not only the operation of addition, but also the special elements 0 and 1 of W, and the successor operation S.

Law of the additive identity element. For every x in W we have $x + 0 = x$ and $0 + x = x$.

We express this fact by saying that 0 is an identity element for the operation $+$.

Law of addition and successor. For every x in W we have $S(x) = x + 1$.

This simply expresses the fact that the addition of 1 to any whole number brings us to the next number in the natural ordering of W.

The preceding two laws, as well as the commutative and associative laws above, all have in common a certain simple form. Namely, each law is expressed by means of a single equation, involving one or more variables, preceded by a phrase such as "For every," or "For all." A law having this form is called an equational identity. There are, however, laws of a more complicated form, such as the following.

Cancellation law for $+$. If x,y,z are any whole numbers such that $x + z = y + z$, then also $x = y$.

The name of this law derives from the fact that the second equation appearing in it may be obtained from the first equation by "cancelling" an occurrence of the letter "z" from both sides. Of course there is no general logical principle which allows such cancellation for an arbitrary operation; the truth of the law for addition rests upon the definition of this particular operation.

The cancellation law for $+$ can be expressed in the following logically equivalent form: If x, y, z are any whole numbers such that $x \neq y$, then also $x + z \neq y + z$. In this form the cancellation law can be seen to be related to the following law connecting the operation $+$ with the relation $<$.

The law of addition over inequality: If x, y, z are any whole numbers such that $x < y$, then also $x + z < y + z$.

A similar law connects $+$ with $\leq$.

Still another law connecting addition with the relation $\leq$ is the following: *For any whole numbers x and y we have $x \leq y$ if, and only if, there is some whole number z such that $x + z = y$*. It is worth noting that this law is used as a *definition* of the relation $\leq$ in some treatments of the theory of whole numbers, where the operation $+$ is introduced before a study of the relations $<$ and $\leq$. In such a presentation of the subject we may follow this definition of $\leq$ by defining $x < y$ to hold if, and only if, x and y are whole numbers for which we have $x \leq y$ *and* $x \neq y$.

We have now examined several laws (i.e., general statements) involving addition, and we wish to turn to the *use* of some of these laws in establishing particular statements about addition. Another way of expressing what we are going to do is to say that we shall use our laws of addition to compute certain sums -- starting from other sums.

To take a very simple example, consider the following partially filled-out addition table:

+	0	1	2	3
0	0	1	2	3
1		2	3	4
2			4	5
3				6

Using only the information contained in the completed part of the table, together with the commutative law of addition, we can obtain all of the information needed to complete the table -- without ever having to go back to a definition of +. We illustrate this by the following theorem, in which the first hypothesis is taken from one of the completed entries of the table above, while the conclusion gives the information needed to provide one of the omitted entries.

Theorem. Using the hypotheses

(i) $1 + 2 = 3$, and

(ii) the commutative law for +,

we may obtain the conclusion

$$2 + 1 = 3.$$

The proof is exceedingly simple.

1. $2 + 1 = 1 + 2$; by logic from hypothesis (ii).
2. $1 + 2 = 3$; by hypothesis (i).
3. $2 + 1 = 3$; by lines 1 and 2 and the logic of $=$.

In explanation of the first line of the proof, we recall that hypothesis (ii), the commutative law for $+$, asserts that $x + y = y + x$ for any whole numbers x and y. Hence in particular we may take x to be 2 and y to be 1, getting $2 + 1 = 1 + 2$ as asserted.

The use of the associative law to obtain certain sums from others requires us to use proofs of somewhat more interest. Consider, for example, the following partially-completed table of values of the addition operation:

Table A

+	1	2	3	4	5	6	7
1	2						
2	3						
3	4						
4	5						
5	6						
6	7						
7	8						

It turns out that the seven entries provided, together with the associative law for $+$, allow us to obtain all the missing entries of the table -- again, without referring to a definition of $+$ at all. Let us illustrate this by obtaining the sum $4 + 3$. The hypotheses (i) below are taken from the completed part of the table; the conclusion allows us to fill in one of the omitted entries.

Theorem. Using as hypotheses

(i) The particular statements,

$1 + 1 = 2$, $2 + 1 = 3$, $4 + 1 = 5$, $5 + 1 = 6$, and $6 + 1 = 7$, and

(ii) the associative law for +,

we may infer the conclusion: $4 + 3 = 7$.

Proof.

1. $2 = 1 + 1$; by hypothesis (i) and logic of = .
2. $4 + 2 = 4 + (1+1)$; from line 1 by logic of = .
3. $=(4+1) + 1$; by hypothesis (ii)
4. $= 5 + 1$; by hypothesis (i) and logic of =
5. $= 6$; by hypothesis (i) and logic of =.
6. $4 + 2 = 6$; by lines 2-5 and logic of = .
7. $3 = 2 + 1$; by hypothesis (i) and logic of = .
8. $4 +3 = 4 + (2+1)$; by line 7 and logic of =
9. $= (4+2) + 1$; by hypothesis (ii)
10. $= 6 + 1$; by line 6 and logic of =
11. $= 7$; by hypothesis (i) and logic of = .
12. $4 + 3 = 7$; by lines 8-11 and logic of = .

After we have studied certain additional laws involving multiplication as well as addition, we shall see that proofs similar to the one above, but involving these other laws, comprise the justification of the algorithms usually taught in elementary school for computing sums and products.

In addition to the laws we have studied above one sometimes finds in books the following:

Closure law for + : If x,y are any elements of W, then x + y is also in W.

The property of + expressed by this law is part of what we mean by saying that + is a two-place operation on W.

There are other sets A, subsets of W, such that for any numbers x,y in A we have x + y in A. Such a set is said to be closed under +. The set P of positive whole numbers, i.e., the set of all whole numbers x such that x > 0, is an example of a set which is closed under +. The set of all even whole numbers is another example.

§5 (Track B)

1. Exercise. In (i)-(vi) below particular examples of either the commutative law for +, the associative law for +, or the law of the additive identity element are given. In each case state which law is being represented and give an additional particular example of that law.

(i) $3 + 0 = 0 + 3$

(ii) $(2 + 0) + 0 = 2 + (0 + 0)$

(iii) $0 + 0 = 0$

(iv) $108 + (2 + 17) = (2 + 17) + 108$

(v) $(2 + 3) + 0 = 2 + 3$

(vi) $(2 + 5) + (3 + 0) = 2 + (5 + (3 + 0))$

2. Class Discussion. At the end of §4, it was stated that the set of all even whole numbers is closed under +. Discuss what is meant by "even whole number", and how the

general concept of an even whole number can be described mathematically. Subsequently have the class present an argument for the fact that the even whole numbers are closed under $+$. Now consider the set of all odd whole numbers in the light of the above discussion. Is this set closed? Make up other sets.

3. Exercise. If x, y and z are any whole numbers, explain why $x + (y + z)$ is a whole number. Suppose also that v is any element of W. Explain why $(x + (y + z)) + v$ is a whole number. How far can we "extend" the above sums -- that is, how many whole numbers (addends) may we sum together and still be assured that our result will be a whole number? Explain.

4. Exercise. Suppose x, y and z are any whole numbers. What supporting statements are needed to justify each of the following: (a) If $x = z$ then $x + y = y + z$

(b) If $x + y = y + z$ then $x = z$

5. Class Discussion. The law of "the" additive identity element tells us that 0 is an additive identity element for W. Discuss the possibility that there is another identity element, say a, for the set of whole numbers under $+$. Have the class collaborate on a proof that in fact, if a is any additive identity element then $a = 0$; explain why, as a result, we are justified in talking about "the" (unique) additive identity element for W.

6. Exercise. Let w, x, y, z be any four whole numbers in this order. (a) Using this order, in how many distinct

ways may these numbers be summed? (See the discussion of the associative law in §4.) (b) $(w + (x + y)) + z$ and $w + (x + (y + z))$ are two of these distinct ways. Using the associative law for + repeatedly, in a step by step manner, show that these expressions denote the same number.

7. Class Discussion. Suppose r, s, t are any whole numbers. What are we talking about when we say that the associative law for + in W gives meaning to the expression "$r + s + t$" ?

8. Exercise. Closely connected with the two-place operation of addition, there are infinitely many one-place operations on W. An example will make this clearer.

Let us denote by S_5 the one-place operation such that, for any whole number x, when S_5 operates on x the resulting whole number is $x + 5$.* We use the notation "$S_5(x)$" because it gives us the same result as applying the one-place operator S five times -- that is, $S_5(x) = S(S(S(S(S(x)))))$. Let us call S_5 the plus-five operator. Analogously, we can define a plus-two, plus-three, or plus-y operator (y being some whole number).

(a) Compute $S_5(0)$, $S_2(4)$, $S_0(5)$.

(b) Formulate a precise definition of the plus-two operator.

(c) Solve the following equations; that is, determine what whole number x must be in order to make the equations true:

(i) $S_5(x) = 7$

(ii) $S_3(x) = 4$

(iii) $S_x(2) = 6$

(iv) $S_x(S_2(3)) = 8$

(v) $S_2(S_x(x)) = 10$

(vi) $S_x(S_4(x)) = 16$.

---

* In symbols, $S_5(x) = x + 5$.

9. Exercise. As was mentioned in §4, some texts define the relation $\leq$ in terms of addition -- that is, $x \leq y$ is defined to hold if, and only if, there is some whole number z such that $x + z = y$. The relation $<$ can then be defined to hold between two whole numbers x and y if, and only if, both $x \leq y$ and $x \neq y$. Starting from the trichotomy law for the relation $<$, and using the above connections between $<$ and $\leq$, and between $\leq$ and $+$, obtain a general statement about the operation $+$.

10. Exercise. In §4 the associative law is used in order to find the sum $4 + 3$, an entry in Table A. Review the procedure involved in computing $4 + 3$ and then using the same hypotheses show that $1 + 3 = 4$.

11. Exercise. Suppose that a, b, c, d, e are five distinct objects and that $F = \{a,b,c,d,e\}$. A two-place operation on F is determined as soon as we have a table indicating which element of F is assigned as the value of the operation when it acts on any given ordered pair of elements of F. Below is a partially completed table for a certain two-place operation $\odot$ on F. Because there are some blank boxes this table does not, by itself, determine the operation $\odot$.

(a) Write out a statement of the commutative law for the operation $\odot$. Assuming that this law holds for $\odot$, complete the partially filled table of values for $\odot$ below.

(b) Describe a kind of symmetry which can be observed in the completed table. Suppose the entries in the table are rearranged by putting the elements of F in the order (a,b,c,d,e), instead of the present order (b,d,e,a,c), in both top and left margins of the table (and then rearranging the entries in the body of the table accordingly). Will the resulting table be symmetric? What if we rearrange the top margin but not the left margin?

(b) Does there exist an identity element in F with respect to the operation $\odot$? Explain.

(c) Compute the elements $(b \odot d) \odot c$ and $b \odot (d \odot c)$ of F. How would you go about establishing the associative law for the operation $\odot$? How many equations would have to be verified by computation?

⊙	b	d	e	a	c
b	c	e	a	b	d
d		b	c	d	a
e			d	e	b
a				a	c
c					e

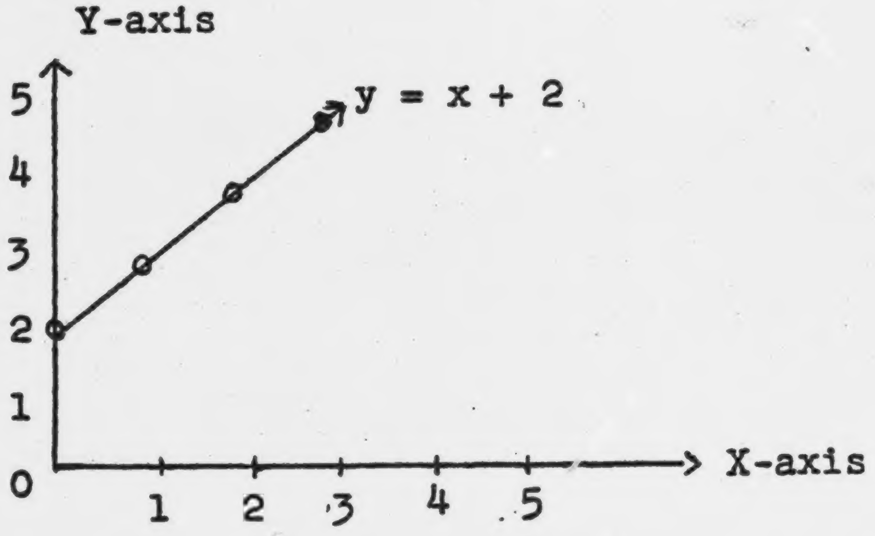
§6 (Track A)

Now let us see what sort of pictures -- or geometric models, to use a more mathematical-sounding phrase -- we can construct to illustrate the concept of addition.

Let us begin with graphs. Consider the equation $y = x + 2$, for example. Its graph consists of all points labeled with ordered pairs of whole numbers (a,b) which satisfy the equation $y = x + 2$ when the first number, a, is substituted for the letter "x" and the second number, b, is substituted for the letter "y". For example, the point (1,3) is on the graph since $3 = 1 + 2$ but (3,1) is not on the graph since $1 \neq 3 + 2$. In order to find several points on the graph we often make a table with two columns headed "x" and "y"; we put several whole numbers in the x-column, and opposite each of these we put a number in the y-column obtained by computation from the equation $y = x + 2$.

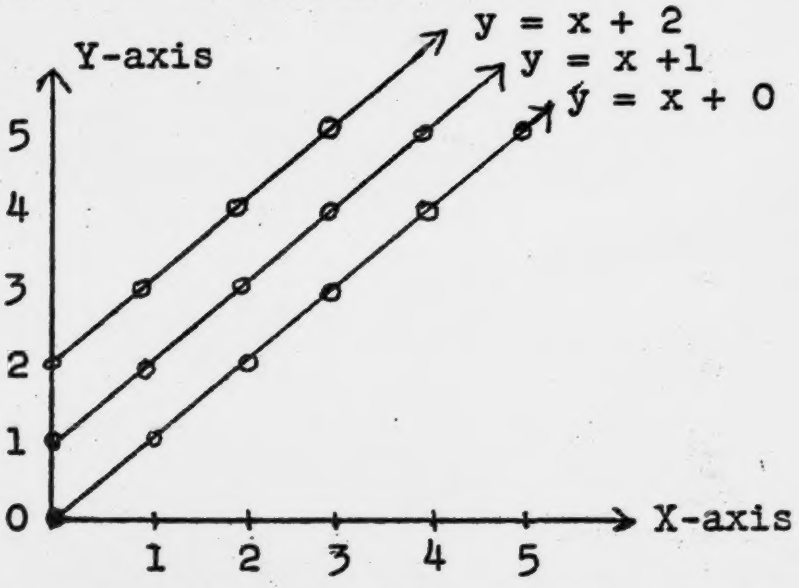
x	y
0	2
1	3
2	4
3	5

Thus the last line of this table tells us that the point (3,5) is on the graph. Our picture of the graph, derived from this table, looks as follows:



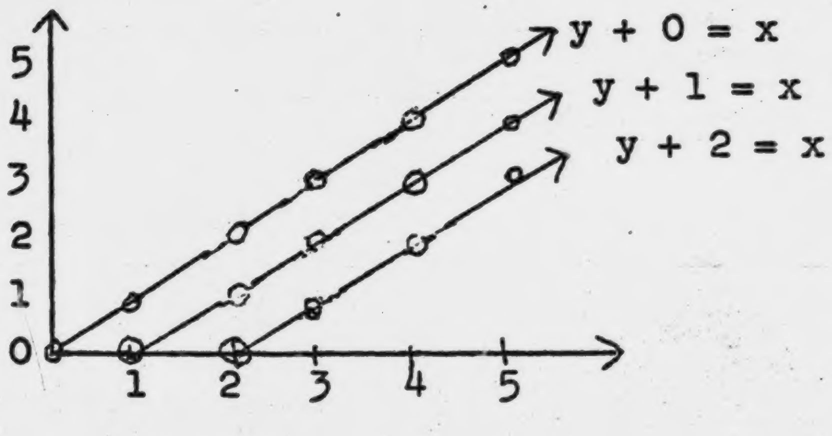
We see that the points of the graph all seem to lie on one straight line. We have placed an arrow at the end of the line joining the points of the graph to remind ourselves that only a limited portion of the graph is shown in the picture; the full graph extends indefinitely in the direction of the arrow.

It is instructive to plot the graphs of several equations, say $y = x + 0$, $y = x + 1$, and $y = x + 3$, on the same picture. Here is the result:

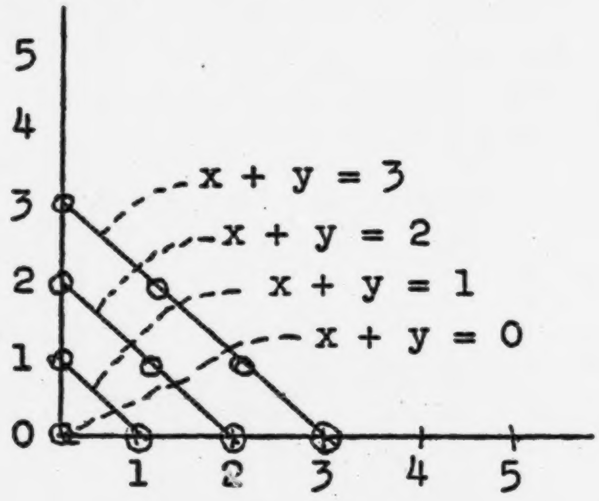


We see that the lines joining the points of these graphs are parallel.

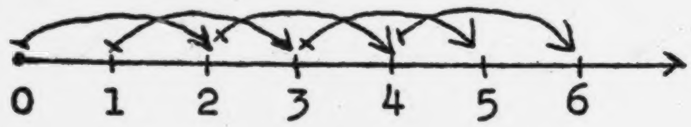
Other pictures connected with the operation + can be obtained by plotting graphs of the equations $y + 0 = x$, $y + 1 = x$, and $y + 2 = x$.



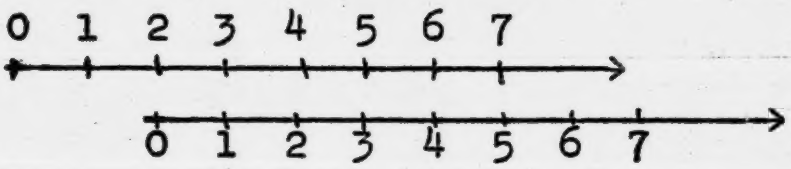
Finally, let us plot the graphs of $x + y = 0$, $x + y = 1$, $x + y = 2$, and $x + y = 3$. Unlike the preceding graphs, these will not extend indefinitely. In fact, there is only one pair of whole numbers which satisfies the equation $x + y = 0$, namely, $(0,0)$, so the graph consists of just one point in this case. Similarly, the equation $x + y = 2$ is satisfied by the ordered pairs $(0,2)$, $(1,1)$, and $(2,0)$, and by no other pair of whole numbers, so the graph contains just three points in this case.



Graphs are not the only kind of pictures we can associate with the operation +. Let us consider the equation $y = 2 + x$. We can get a picture for it by drawing a number line and a series of arrows: An arrow is to start from each numbered point, and to end at another point whose number is obtained from the first number by adding it to 2.



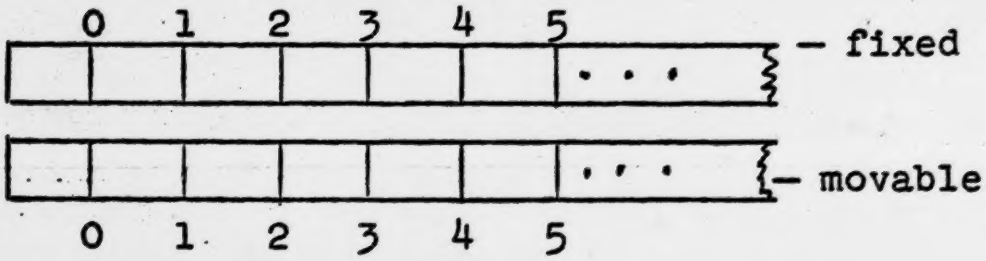
These arrows suggest a motion of the number line, in which the line moves 2 units to the right. (We thus get a "motion picture" for the equation $y = 2 + x$!) If we indicate this motion by drawing the initial position above and the final position below, the result is as follows:



We see that if we choose a number from the lower line and substitute it for the letter "x" in the equation $y = 2 + x$, the corresponding value of the letter "y" will lie on the upper line directly above the initial number. Of course, similar equations, such as $y = 1 + x$ or $y = 5 + x$, would correspond to motions of the number line 1 unit to the right, or 5 units to the right, respectively.

The geometric ideas we have just considered lead to a mechanical device known as the slide rule. This consists of two strips of wood or other material, each imprinted with a scale of numbers, so arranged that one strip can slide alongside the other. Slide rules are commercially manufactured and widely employed, especially by engineers; these slide rules generally employ logarithmic scales and are used as an aid in multiplication. However, by employing scales such as we use on a number line, we can construct a slide rule which enables us to find sums in a mechanical manner.

We consider a slide rule consisting of fixed and movable strips of wood placed alongside each other, on each of which a number line has been imprinted. We will picture the fixed strip above, the movable strip below.

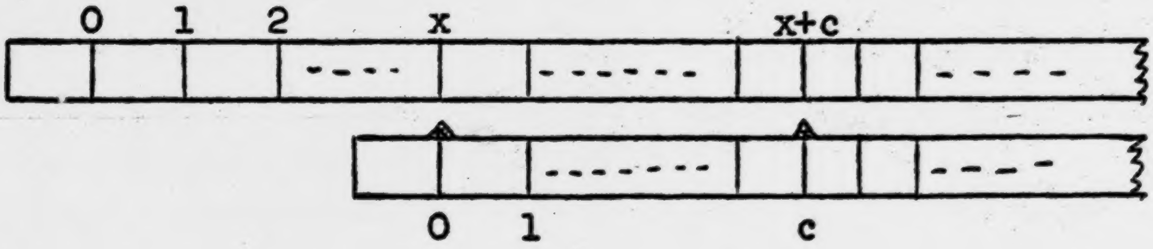


Now the use of this device to compute any sum $c + x$ of given whole numbers c and x, is as follows:

(i) We move the 0 point on the lower scale until it is opposite the number c on the upper scale.

(ii) We find the point x on the lower scale and read off the number y on the upper scale which is opposite it. This number y is the desired sum $c + x$.

The fact that we get $y = c + x$, as claimed in (ii) above, may be seen from the consideration of the moving number line which precedes our discussion of the slide rule.



§7 (Track B)

1. **Exercise.** On one set of coordinate axes, graph the following equations: (a) $x + y = 4$

(b) $x + 2 = y$

(i) Determine the set A of all lattice points whose coordinates satisfy equation (a).

(ii) Let B be the set of all lattice points whose coordinates satisfy equation (b). Do all of these points appear on the part of the graph you have pictured? Explain.

(iii) Find $A \cap B$.

(iv) How is the answer to (iii) related to your picture of the graphs of equations (a) and (b)?

2. Class discussion. A teacher should have practice in making up "word problems" for presentation to pupils for the purpose of illustrating mathematical ideas. Ask the class to find word problems which can be represented by equations (a) and (b) in Exercise 1 above.

A "classical example" of such a problem would involve the ages of children. Example: The ages of John and Jane now add up to 4, and in two years from now John will be as old as Jane is now; how old is each child? For another example, we may seek to determine the lengths x and y of two sticks of unknown length, in a situation where we have on hand sticks of known length 2 and 4.

Discuss the solutions of these problems in terms of graphs. Find an additional pair of equations whose graphs intersect, and "fit" a few word problems to these graphs.

3. Exercise. (i) Let p be the horizontal line segment whose endpoints have coordinates $(2,1)$ and $(6,1)$ -- we shall express this by writing $p = [(2,1), (6,1)]$. Also, let q be the horizontal line segment $[(3,5), (9,5)]$. Draw a pair of coordinate axes and sketch the two segments p and q. Mark the approximate midpoints (on the drawing

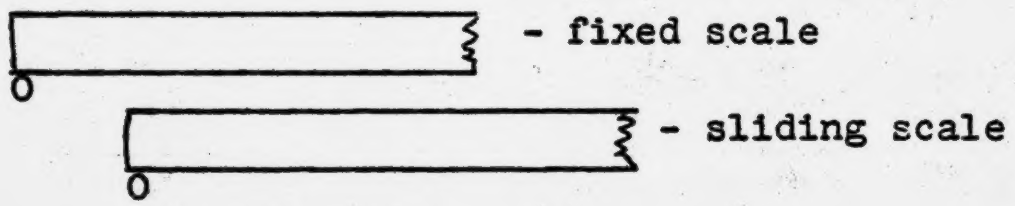
of these segments) by eye -- without measuring -- and estimate the coordinates of these midpoints by visual inspection.

(ii) On another graph draw the vertical line segments [(1,1), (1,3)] and [(2,2), (2,6)]. Assign letters as names for these segments, and estimate the coordinates of the midpoints of each of them, again by visual inspection.

(iii) Study the coordinates for the four midpoints supplied in (i) and (ii) above. Can you find a relationship which holds in each case between the coordinates of the midpoint and the coordinates of the endpoints?

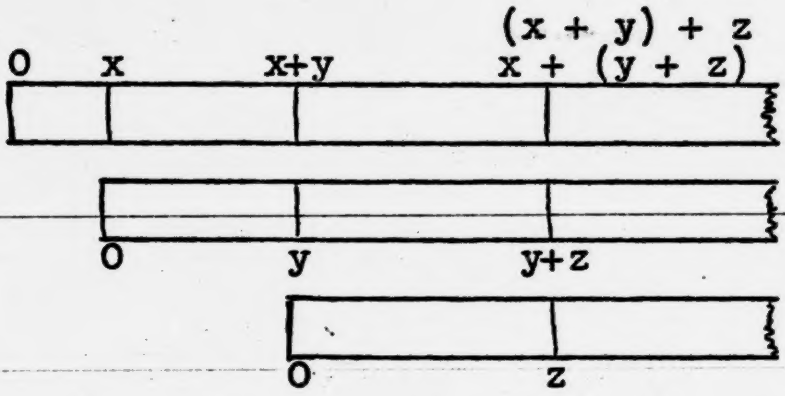
(iv) Using the relationship found in (iii) above, what would be the coordinates of the midpoint of the segment [(2,1), (8,3)]? Make a drawing and check your answer.

4. Exercise. The sliding scale of the slide rule below has been moved into position to compute the sum 3 + 4. Put in the necessary numerals on both scales to enable us to compute 3 + 4, and circle the numeral represented by this sum on the slide rule.



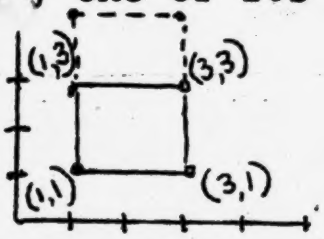
5. Class discussion. Discuss the feasibility of actually constructing some sort of a slide rule to be used in an elementary school -- size, material, details. Consider the possibility of using a "super slide rule" consisting of three scales each of which can be moved independently while the other

two are held fixed relative to each other. Such a super slide rule can be used to illustrate the associative law of addition: discuss how this can be done.

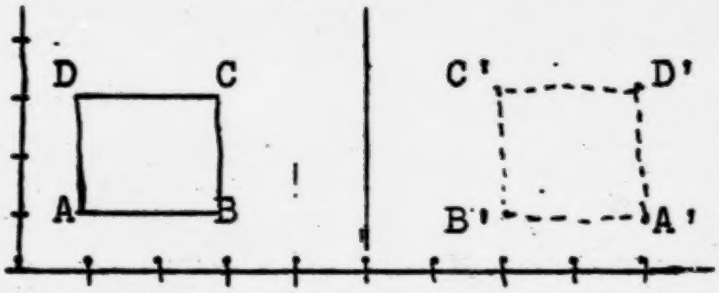


§8 (Track C)

1. A moving game. The following "game" is suggested as a way to give children more practice with coordinates while simultaneously enabling them to develop relevant geometric intuition. The idea is to have them move around a simple figure, such as a square or a triangle, in a patterned way upon a system of coordinate axes. For example, the teacher connects four lattice points to form a square, asking the children to determine the coordinates of its vertices. Next the children are asked to determine the coordinates of the vertices of the square obtained by "flipping" the given square over one of its sides, as below. (A cutout cardboard square of the same size can be placed on the chalkboard and flipped over to demonstrate what is meant. Of course, the square can be flipped over about any one of its four sides.



As the children gain in visualizing geometric motions a horizontal line can be drawn at some distance from a square, and the question can be put as to where the square would land if it were flipped over this line in a kind of mirror-image projection. The vertices of the given square can be labeled A, B, C, D, and then those of the square in the new position should be labeled A', B', C', D' in such a way that A' shows the new position, after flipping, of the original vertex A; and similarly for each vertex. Finally, when coordinates of the vertices of the original square are given, the coordinates of the vertices of the square in the new position should be computed.



§9 (Track A)

Finally, let us turn our attention to the principle of mathematical induction which we have mentioned several times without formulating it explicitly. While its precise formulation will seem rather more complex than the general statements we have considered heretofore, there is really a very simple idea which lies behind it. In fact, this principle is nothing more than a precise way of saying that all whole numbers are obtained by starting with 0 and applying the successor operation, S, over and over again.

Suppose that A is a set whose elements are whole numbers, i.e., $A \subseteq W$. We are going to suppose two things about A: First, that the number 0 is an element of A, and second, that A is closed under the operation S. This means that whenever S is applied to an element of A, the resulting number is also in A. Now let us combine these two suppositions and see what follows.

Since 0 is in A we can apply S to it obtaining 1; thus 1 is in A (since A is closed under S). Now, applying S to 1 we get 2; thus 2 is in A (since A is closed under S). Now, applying S to 2 we get 3; thus 3 is in A (since S is closed under S). Continuing in this way we see intuitively that *every* whole number is in A -- since every such number can be reached by successive applications of the operation S to 0. The fact that this conclusion can be drawn from our two assumptions about A is precisely the content of the principle of mathematical induction. Let us now formulate this explicitly.

Principle of mathematical induction. *Let A be any set of whole numbers such that*

(i) *0 is in A, and*

(ii) *Whenever a whole number x is in A then also S(x) is in A.*

Then every whole number must be in A.

This principle can be used to prove a great many facts about the whole number system. For this reason, when mathematicians treat the number system as an axiomatic theory, they often include the principle of mathematical induction as one of the axioms. Most people know that the Greek mathematician Euclid developed geometry as an axiomatic theory, but the reasons which impelled him to do so are applicable to every branch of mathematics. It was an Italian mathematician, Peano, who first set up an axiomatic theory of whole numbers, around 1890. Nowadays mathematicians treat all parts of mathematics from the axiomatic viewpoint.

The desirability of setting up axioms for a theory arises from the simple recognition that every proof reaches its conclusion only after starting from some assumptions. If we go back and try to prove those assumptions, we must start those proofs from other assumptions. If we are not to be led by this process into circular reasoning, we must decide to start somewhere with propositions which we do not try to prove. These are the axioms of our system. Of course there is nothing intrinsically umprovable about these axioms: One can always find a new set of axioms and use them to prove the propositions taken as axioms in the first system.

Among the axioms used by Peano were the principle of mathematical induction, and the following two general statements connecting + with 0 and with S respectively.

Axiom 1. For every x in W we have
$$x + 0 = x.$$

Axiom 2. For every x, y in W we have
$$x + S(y) = S(x + y).$$

Let us see how the associative law for $+$ can be obtained as a theorem in this system.

Theorem. Let us assume Axioms 1, 2 and the principle of mathematical induction. Then for all x, y, z in W we have
$$(x + y) + z = x + (y + z).$$

Proof.

1. Let x, y be any whole numbers.
2. Having chosen x and y, let us form the set A of all those whole numbers z (if any) for which it is true that $(x + y) + z = x + (y + z)$.
3. We claim that 0 is in this set A. For $(x + y) + 0 = x + y$, by Axiom 1. And $x + (y + 0) = x + y$, because $y + 0 = y$ by Axiom 1. Combining these two equations we get $(x + y) + 0 = x + (y + 0)$, which means that 0 is in A by definition of A (Step 2, above).
4. Now suppose that we choose any number z from our set A. We claim that we must also have $S(z)$ in A. To see this, let us compute.

(a) $(x + y) + z = x + (y + z)$, since z was chosen from A (Step 4), using the definition of A (Step 2).

(b) $S((x + y) + z) = S(x +(y + z))$, by logic from (a).

(c) $S((x + y) + z) = (x + y) + S(z)$, by Axiom 2. (The x of Axiom 2 is taken to be $x + y$, and the y of Axiom 2 is taken to be z.)

(d) $(x + y) + S(z) = S(x + (y + z))$, by logic, (b), and (c).

(e) $S(x + (y + z)) = x + S(y + z)$. By Axiom 2.

(f) $(x + y) + S(z) = x + S(y + z)$. By logic, (d) and (e).

(g) $S(y + z) = y + S(z)$. By Axiom 2.

(h) $x + S(y + z) = x + (y + S(z))$. By logic and (g).

(i) $(x + y) + S(z) = x + (y + S(z))$. By logic,(f), and (h).

(j) $S(z)$ is in the set A. By (i) and the definition of A (Step 2).

5. Every whole number is in the set A. For 0 is in A (Step 3), and whenever a whole number z is in A we have also $S(z)$ in A (Step 4). Hence we may apply the principle of mathematical induction to conclude that all whole numbers are in A.

6. For every whole number z we have $(x + y) + z = x + (y + z)$. From Step 5 by the definition of set A (Step 2).

7. Since x, y were any whole numbers (Step 1) the statement of Step 6 gives the desired conclusion of our theorem.

§10 (Track B)

1. Exercise. Assume Axioms 1 and 2 and the principle of mathematical induction, as given in §9. Using these, prove that for all x in W we have

$$x + 0 = 0 + x$$

(a special case of the commutative law for $+$). Hint: Form the set A consisting of all those numbers x of W -- if any -- for which we do, in fact, have $x + 0 = 0 + x$. Then apply the principle of mathematical induction to this set A by using our assumptions to show (i) 0 is in A, and (ii) whenever a number y is in A, then the number $S(y)$ must also be in A.

2. Comment. Whenever the principle of mathematical induction is employed in a proof, the application must begin by defining a certain set A of whole numbers. A successful application of the principle will result in the conclusion that all numbers are in this set, and from this information we must be able to establish our desired result. It is for this reason that in Exercise 1 above we chose A to be the set of those numbers x for which it is true that $x + 0 = 0 + x$.

In many problems, such as the one just considered, there is only one reasonable way in which to define the set A to which mathematical induction is to be applied. In other cases, however, several possible definitions suggest themselves -- one or more of these may work, but others may not. A case in point is to be found in §9, where the principle of mathematical induction was employed to prove the associative law of addition. Let us review the circumstances.

The associative law for $+$ states that for all whole numbers x, y, z we have $x + (y + z) = (x + y) + z$. In our proof we first chose any whole numbers x and y. Then, having chosen and fixed these two numbers, we tested each whole number z to see whether or not $x + (y + z) = (x + y) + z$, and we formed the set A of all those numbers z for which the equation holds. We were able to show by mathematical induction that all whole numbers are in this set A, which established the associative law as desired.

Instead of the set A defined above, it is perfectly natural to consider a set B defined in a different way, as follows. We first choose any whole numbers y and z. Then, having chosen and fixed these two numbers, we test each whole number x to see whether or not $x + (y + z) = (x + y) + z$, and we form the set B of all those numbers x for which this equation holds. Now, if we could use mathematical induction to show that all whole numbers are in this set B, we could again conclude that the associative law for $+$ is true. But when we try to apply mathematical induction to B, we get stuck: There is difficulty, for example, in showing that 0 is in B, which must be done in order to apply mathematical induction.

Still another set, C, is a natural one to consider in trying to prove the associative law for $+$. (How would we define C?) But we get stuck in trying to prove 0 is in C, just as we did for the case of B.

How can we tell in advance that the set A will work, while the sets B and C will not? There is no general way, except trial and error. Thus whenever an application of the principle of mathematical induction is attempted, one has to experiment with the definition of the set A to which the principle will be applied.

3. Exercise. In Exercise 1 above a special case of the commutative law for + was established, using Axioms 1 and 2 and the principle of mathematical induction (as formulated in §9). Now use those same assumptions to prove the commutative law for + in full generality. (Hint: of course you can use the result of Exercise 1 as a part of your proof.)

Chapter 4. Multiplication

§1 (Track A)

In Chapter 3 we presented two alternative definitions of addition, and in this chapter we shall deal in the same way with the operation of multiplication.

First Definition of Multiplication. We define a two-place operation called multiplication, and symbolized $\cdot$, on the set W of all whole numbers. If we operated with $\cdot$ on any given whole numbers x and y, the resulting number, $x \cdot y$, is called the product of x and y, and is obtained by the formula

$$x \cdot y = \underbrace{y + y + \ldots + y}_{x \text{ occurrences of } y} ,$$

that is, by repeated addition of the number y a total of x times. In case x is 0 we do not add any occurrences of the number y and we define

$$0 \cdot y = 0 .$$

This definition of multiplication as repeated addition is, of course, familiar as the most common method of introducing multiplication in the elementary schools. The need for a special clause in the definition covering the case where the first factor is 0 is sometimes overlooked; of course, it is unnecessary in schools which first study the system of counting numbers, 1, 2, 3, ..., where 0 is not present.

However, one may ask why we choose to define $0 \cdot y$ as 0, rather than as y, say, once it is recognized that the main clause of the definition of $x \cdot y$ does not give a clear and precise result for the case where x is 0. There are several ways in which this question may be answered.

1. From the viewpoint of applications: We know that if x is the number of boys in a class, and if y is the number of marbles each boy has, then $x \cdot y$ is the total number of marbles in the class (assuming that neither teachers nor girls have any marbles). Now for the case where x is 0 -- an all-girls' class -- there will, in fact, be no marbles, so we do want to have $0 \cdot y = 0$ for this application.

2. From the point of view of simple laws: As we have learned in elementary school, and as we shall see below, the operation of multiplication satisfies a commutative law. That is, $x \cdot y = y \cdot x$ for any whole numbers x and y. Now if y is any of the counting numbers we easily see that $y \cdot 0 = 0$. For instance, using our definition of $\cdot$ above, we compute $1 \cdot 0 = 0$, $2 \cdot 0 = 0 + 0 = 0$, $3 \cdot 0 = 0 + 0 + 0 = 0$, etc. Hence, we must make our definition of multiplication give a value of 0 to the products $0 \cdot 1$, $0 \cdot 2$, $0 \cdot 3$, etc., otherwise the commutative law would fail.

3. Compatibility with other definitions: We shall give below an alternative definition of multiplication which does not require any special clause for the case where the first given number is 0. The original definition gives the same value for $x \cdot y$ as the new definition for all whole numbers

x and y where $x \neq 0$, so it is natural to design the special clause of the original definition in such a way that the products $0 \cdot y$ also have the same value as under the new definition. This value is 0.

So much for the various reasons which motivate the special clause of our first definition of multiplication. Before we can proceed, however, we must still examine the main clause of our definition. If we employ this to obtain the product $4 \cdot 2$, for example, we find that we must compute the sum $2 + 2 + 2 + 2$. But what does this mean? Since addition was defined to be a two-place operation we can only add two numbers at a time, yet here we seem to be adding four numbers.

The preliminary answer to this question is that we must, indeed, add only two numbers at a time, and so the four given numbers cannot be added simultaneously but should be taken in some order so that we carry out a succession of additions, each time adding just two numbers. To indicate the order of carrying out the additions we use parentheses. For example, one order would be $((2 + 2) +2) +2$. But this is not the only possibility. Others are $2 + (2 + (2 + 2))$ and $(2 + 2) + (2 + 2)$, and there are still two other possibilities. How do we know that these five orders of summation will lead to the same result? If they do not, which of the resulting sums is meant by the notation $2 + 2 + 2 + 2$?

It turns out that all orders of summation give the same result, 8, and so the notation "$2 + 2 + 2 + 2$" may be used to refer unambiguously to this number. This can be proved by several applications of the associative law for +. For example, to show that

$$((2 + 2) + 2) + 2 = 2 + (2 + (2 + 2))$$

we first get

$$(*) \qquad ((2 + 2) + 2) + 2 = (2 + 2) + (2 + 2)$$

by taking x to be $2 + 2$, y to be 2, and z to be 2 in the equation

$$(x + y) + z = x + (y + z)$$

which is involved in the associative law. Then we get

$$(**) \qquad (2 + 2) + (2 + 2) = 2 + (2 + (2 + 2))$$

by taking x to be 2, y to be 2, and z to be $2 + 2$ in the equation of the associative law. And finally we combine equations (*) and (**) by the logic of equality to get the desired result.

Similarly, whenever we add more than two whole numbers in a given order, repeated use of the associative law will show that any order for computing the sum by a succession of additions of two numbers at a time, leads to the same result as any other such order. (A single proof covering an arbitrary number of terms to be added can be given, but requires the principle of mathematical induction.) It is only because of this result that we are permitted to use the notation $y + y + \ldots + y$ in our definition of multiplication without providing parentheses to indicate some particular order for carrying out the additions. Thus this definition of multiplication requires justification

which is usually taken for granted without mention in the elementary schools. (Compare Exercises 6, 7 of

Let us now turn to the alternative definition of multiplication to which we have alluded above. We have already indicated that our first definition of multiplication resembles the second definition of addition given in Chapter 3, insofar as the defined operation is expressed in terms of repeated application of some other operation introduced earlier. It is thus natural to inquire whether we can now give another definition of multiplication which resembles, in form, the first definition of addition in Chapter 3. In order to define $x \cdot y$ by this method, where x and y are any whole numbers, we would first choose sets A and B such that $n(A) = x$ and $n(B) = y$, i.e., such that the number of elements in A is x and the number in B is y; we would then combine A and B somehow to obtain a new set, C; and finally we would count the number of elements in C and declare that the resulting number, $n(C)$, is the value of $x \cdot y$. This is, indeed, what we shall do. But how are we to combine the sets A and B to obtain C? Taking the union $A \cup B$ will not do as part of a definition of multiplication, since that method leads to the operation of addition (providing A and B are disjoint sets). We need another method of proceeding from the given sets A and B to the new set, C, and we now turn to a description of this.

Definition. If A and B are any two sets, then the cartesian product of A and B, denoted $A \times B$, is another set, whose elements are all of the ordered pairs (x,y) which can be formed using any element x of A as its first component and any element y of B as its second component.

The adjective "cartesian" is taken from the name of Descartes, a famous French mathematician and philosopher of the 18th century, whose invention of analytic geometry brought arithmetic and algebra into a close relation to Euclidean geometry. Descartes' use of ordered pairs of numbers (x,y) as coordinates for points in a plane has already been encountered in our study of graphs (Chapters 2, 3).

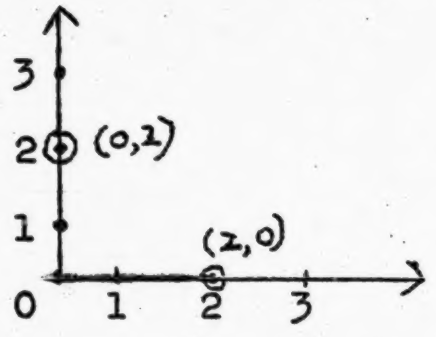
As a simple example of the cartesian product of sets, let

$$A = \{0, 3, 4\} \quad \text{and}$$
$$B = \{2, 3\}.$$

Then we have

$$A \times B = \{(0,2), (0,3), (3,2), (3,3), (4,2), (4,3)\} \quad \text{and}$$
$$B \times A = \{(2,0), (2,3), (2,4), (3,0), (3,3), (3,4)\} .$$

Notice that $A \times B \neq B \times A$ because the ordered pair $(0,2)$ is an element of $A \times B$ but not of $B \times A$. It is true that the ordered pair $(2,0)$ is an element of $B \times A$, but $(2,0) \neq (0,2)$. The fact that the ordered pairs $(2,0)$ and $(0,2)$ are not the same is readily perceived by noticing that they are coordinates for two quite different points with respect to a pair of coordinate axes.



It is worth noting that even though $A \times B \neq B \times A$ in this example, we do have $n(A \times B) = n(B \times A)$.

Now we are ready for the

Second Definition of Multiplication. We define multiplication to be the two-place operation $\cdot$ on the set W such that, if x and y are any whole numbers then the product $x \cdot y$ is obtained by choosing any sets A and B such that $n(A) = x$ and $n(B) = y$, and then setting $x \cdot y = n(A \times B)$.

We have already had several examples of definitions which require justification of one or another sort, and the present one is no exception. The kind of justification needed here is similar to the one needed in the case of the First Definition of addition. The need arises because if one person chooses sets A and B such that $n(A) = x$ and $n(B) = y$, and if another person chooses different sets C and D such that $n(C) = x$ and $n(D) = y$, then the first person will compute $x \cdot y$ to be $n(A \times B)$ and the second will compute $x \cdot y$ to be $n(C \times D)$ -- but how do we know that $n(A \times B)$ will be the same number as $n(C \times D)$? Clearly we need a theorem.

Theorem. If A, B, C, and D are any sets such that $n(A) = n(C)$ and $n(B) = n(D)$, then we will also have $n(A \times B) = n(C \times D)$.

A proof of this theorem requires a heavy use of the function concept (Chapter 1, §4) which we have not developed in much detail, and so we shall not give such a proof. However, the reader should see that the Second Definition of Multiplication would be unsatisfactory if it were not for the fact expressed by this theorem. Incidentally, a comparison of the statements of this theorem and the corresponding theorem justifying the definition of addition will show certain differences of detail which are not essential -- for example, the use of letters "x" and "y" in the earlier theorem could have been eliminated, or such use could have been incorporated in the later theorem. However, there is one important difference: The requirement that the sets A and B be disjoint is essential to the definition of addition and to the theorem justifying it; but there is no corresponding requirement in the definition of multiplication or its justifying theorem.

As we have observed in the case of addition, the reason why we say that the two definitions of this chapter, which seem so dissimilar, both define the same operation, multiplication, is that both definitions lead to the same value of the product $x \cdot y$ for any whole numbers x and y. In a particular case this is easy enough to see. For example, according to the first definition we have $3 \cdot 2 = 2 + 2 + 2$, and so $3 \cdot 2 = 6$. Using the second definition we may choose the sets $A = \{0,3,4\}$ and $B = \{2, 3\}$ used above in illustrating the concept of cartesian product, and by counting the elements of the set $A \times B$

to obtain $n(A \times B)$ we find $x \cdot y = 6$ by this definition too.

But this special case does not give much insight into how ~~we can know that the two definitions will always lead to the~~ same value. The following considerations may help the reader to see this.

Let x and y be any whole numbers whatever, $x \neq 0$, and let us compute $x \cdot y$ by the first definition of multiplication. We get, of course, $x \cdot y = y + y + \ldots + y$, where the term y occurs x number of times on the right. Now, recalling the first definition of addition, we see that to compute this sum $y + y + \ldots + y$ we must choose x number of sets $A_1, A_2, \ldots, A_x$, each having y as the number of its elements, and no two of the sets having an element in common. We must then form the union $A_1 \cup A_2 \cup \ldots \cup A_x$ of all these sets by combining all of their elements into one big set. Finally, the number of elements in this union, $n(A_1 \cup A_2 \cup \ldots \cup A_x)$, will be the desired sum $y + y + \ldots + y$, (having x number of terms), that is we will have $x \cdot y = n(A_1 \cup A_2 \cup \ldots \cup A_x)$.

Now how shall we choose our sets $A_1, A_2, \ldots, A_x$? Since the number of elements in each set must be y, we might think of taking each of these sets to be $\{1, 2, \ldots, y\}$. However, we cannot really take all of the sets $A_1, A_2, \ldots, A_x$ to be the same, for no two of the sets may have an element in common. So let us modify our first idea and take

$$\begin{array}{l} A_1 = \{(1,1),\ (1,2),\ \ldots,\ (1,y)\} \\ A_2 = \{(2,1),\ (2,2),\ \ldots,\ (2,y)\} \\ \vdots \qquad\quad \vdots \qquad\quad \vdots \qquad\qquad\quad \vdots \\ A_x = \{(x,1),\ (x,2),\ \ldots,\ (x,y)\}\ . \end{array}$$

Then clearly each of the sets $A_1, A_2, \ldots, A_x$ will have y as the number of its elements, yet no two of these sets will have an element in common. Thus, our first definition of multiplication (combined with our first definition of addition) gives

$$x \cdot y = n(A_1 \cup A_2 \cup \ldots \cup A_x)\ .$$

But what is this set $A_1 \cup A_2 \cup \ldots \cup A_x$? It consists of all ordered pairs (p, q) where p may be any of the numbers $1, 2, \ldots, x$ and q may be any of the numbers $1, 2, \ldots, y$. In other words, if we set

$$C = \{1, 2, \ldots, x\} \text{ and } D = \{1, 2, \ldots, y\}$$

then $A_1 \cup A_2 \cup \ldots \cup A_x$ is nothing other than the cartesian product $C \times D$. But since clearly $n(C) = x$ and $n(D) = y$, we see that $x \cdot y = n(C \times D)$ by the second definition of multiplication. In this way we see that the two definitions of multiplication lead to the same value for $x \cdot y$, namely $n(A_1 \cup A_2 \cup \ldots \cup A_x)$ or, otherwise written, $n(C \times D)$.

In the reasoning above we considered quite arbitrary whole numbers x and y, except that we assumed $x \neq 0$. It will be recalled that for the case $x = 0$ the first definition of multiplication has a special form. We leave the reader to check that the two definitions of multiplication lead to the same results in this case also.

§2 (Track B)

1. Parentheses and ordered pairs. In §1 we mention the use of parentheses to indicate the order for carrying out a succession of operations, in this case addition. Because addition is defined as a two-place operation, in principle we can only add two numbers at any given time. Thus the expression "2 + 2 + 2 + 2" is ambiguous insofar as it does not indicate in what order the indicated additions are to be performed.

The expression "((2 + 2) + 2) + 2" does indicate precisely an order for carrying out the indicated additions. The rule is that we always begin with the operations indicated with the innermost pairs of parentheses, and continue performing operations "going outward" to evaluate larger and larger portions of the given expression. Thus, to evaluate ((2 + 2) + 2) + 2 we would proceed as follows:

First, 2 + 2 [= 4],
Second, (2+2) + 2 [= 4 + 2] [= 6],
Last, ((2+2) +2) + 2 [= (4+2) + 2] [= 6+2] [= 8] .

On the other hand, to evaluate (2 + 2) + (2 + 2) the procedure would be:

First, 2 + 2 [= 4] -- this combination appears twice in the given expression --
Second, (2+2) + (2+2) [= 4+4] [= 8] .

In this particular case both of the numbers ((2 + 2) + 2) + 2 and (2 + 2) + (2 + 2) turn out to be 8 (as indicated in the text of +1). However, in general the way in which parentheses are distributed to indicate the order of performing operations will affect the numerical value of the result of the computation. (See Exercise 2, below.)

If parentheses are put into an expression in a haphazard manner, the result may be meaningless. For example, the expression "$(2+)(3+5)$" is meaningless since the first pair of parentheses is an innermost pair which does not indicate a pair of numbers to be added. Similarly, the expression $(2+(5+)4)$ is meaningless.

A completely different use of parentheses is involved in forming the name of an ordered pair. If the names of two objects are separated by a comma, we put parentheses around the resulting expression to form a name of the ordered pair having these objects as its first and second elements respectively. This use of parentheses has nothing to do with the order of performing operations.

2. Exercise. (a) The expression "$1 + 2 \cdot 5$" is ambiguous since there is no indication as to the order for performing the indicated operations. Insert parentheses in this expression to indicate one order for carrying out these operations, and then (starting over) insert parentheses to indicate another order. Evaluate the expression in each case.

(b) In the ambiguous expression "$2 + 2 \cdot 2 + 2$", parentheses can be introduced so as to indicate five different orders for performing the indicated operations. If we carry out these five methods of evaluating the expression, how many different numerical values do we get?

3. Exercise. In §1 there is a discussion of the use of the associative law of addition in order to show that $((2 + 2) + 2) + 2 = 2 + (2 + (2 + 2))$. (a) Show this equation to be true using only the commutative law for addition and logical laws.

(b) Note that the associative law can be replaced by the commutative law in this way only in very special cases. Give two more special cases where this is possible.

4. Classroom discussion. Review the line marked (*) in the informal proof, using the associative law of addition, of the equation $((2 + 2) + 2) + 2 = 2 + (2 + (2 + 2))$; that is, clarify the use of the letters "x", "y" and "z". Have the class give additional particular examples of the associative law of addition where sums are substituted for the variables x, y and z.

5. Exercise. One of the definitions of multiplication (the first one given in §1) involves repeated addition; similarly one of the definitions of addition (the second one given in Chapter 3) makes use of repeated applications of the successor operation. (a) Review these two definitions and then combine them to obtain a definition of multiplication directly in terms of repeated use of the successor operation.

(b) Using this new definition of multiplication compute the product $3 \cdot 2$.

6. **Exercise**. (a) For what sets A is it true that $A \times A = A$?

(b) For what sets B is it true that $n(B \times B) = n(B)$?

7. **Exercise**. Let A be the set of whole numbers less than 5. (a) List the elements of $A \times A$ and, by counting, compute $n(A \times A)$.

(b) Let B be the subset of $A \times A$ consisting of all those ordered pairs of $A \times A$ whose first and second members are the same. List the elements of B.

(c) Do the same for the subset C of $A \times A$ consisting of those ordered pairs of $A \times A$ whose second members are equal to twice the value of the first member.

(d) Draw pictures of B and C on one graph.

8. **Exercise**. Recall that the set-theoretic definition of addition (in terms of the union of sets, Chapter 3) required a justifying theorem. So does the set-theoretic definition of multiplication (in terms of the cartesian product of sets, §1). Compare the statements of these two theorems. Reformulate these two theorems, eliminating the use of the letters "x" and "y" in the former and incorporating them in the latter.

9. **Exercise**. At the conclusion of §1 the two definitions for finding the product $x \cdot y$ were shown to be equivalent for all cases except where x is zero. Examine this case by using both definitions of multiplication to compute $0 \cdot y$.

10. Exercise. Suppose $A = \{(1,1), (1,2), (1,3)\}$,
$B = \{(2,1), (2,2), (2,3)\}$,
$C = \{(3,1), (3,2), (3,3)\}$, and
$D = \{(4,1), (4,2), (4,3)\}$.

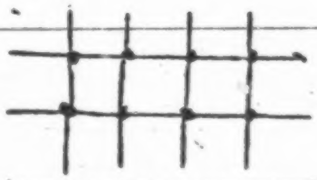
List the elements of sets E and F such that $E \times F = A \cup B \cup C \cup D$.

11. Exercise. Describe in words the elements of the set $W \times W$, where W is the set of all whole numbers. Can a complete list of these elements be put down?

§3 (Track C)

1. Repeated Addition. A number line is helpful for teaching multiplication in the early grades, using the repeated addition approach. For example the product $x \cdot y$ can be obtained by x number of jumps (beginning at 0), each jump being y units in length. A number line painted on the floor would be most useful for the physical jumping involved in computing products by the above method. Analogously, the standard game of Giant Steps could be reformulated to involve products. ["Johnny, you may take 4 jumps each of 2 units length." Johnny answers "Teacher, may I take 8 steps?" "Yes, you may."] It is also possible to use the slide rule described in Chapter 3, moving one stick x number of times to arrive at the point $x \cdot y$ on the fixed part.

2. Cartesian Products. To use this approach to multiplication one need not, of course, mention the name "Cartesian Product." For example, the product $4 \cdot 2$ can be thought of as the number of intersections of 4 vertical columns and 2 horizontal rows as pictured below.



The geo-boards mentioned in Chapter 2, §12 are useful for considering such rectangular arrays of points: In each problem the nails under consideration may be surrounded by a rubber band. For the later grades, Cartesian Products can be introduced without using numbers. For example, if a boy has 3 different colored shirts he can wear, and 2 different colored pants, then the cartesian product of his set of pants with his set of shirts gives the set of the six possible combinations he can wear.

§4 (Track A)

Now that we have seen two alternative definitions for the operation of multiplication, $\cdot$, on the set W of all whole numbers, let us consider some of the laws, or general statements, which hold about it. Two very basic ones have already been encountered in connection with addition.

Commutative Law for Multiplication: If x, y are any whole numbers then $x \cdot y = y \cdot x$.

Among the particular cases included in this law, for example, is $3 \cdot 5 = 5 \cdot 3$. If we seek to verify this by using the first definition of multiplication, we see that $3 \cdot 5 = 5 + 5 + 5$ and $5 \cdot 3 = 3 + 3 + 3 + 3 + 3$, so that what must be shown is that $5 + 5 + 5 = 3 + 3 + 3 + 3 + 3$. By carrying out the indicated sums we can find that this is, indeed, the case. But this method of procedure is not very helpful in seeing that other particular instances of the commutative law will hold. For instance, to verify that $2 \cdot 4 = 4 \cdot 2$ we must compute the sums $4 + 4$ and $2 + 2 + 2 + 2$ and show that they are the same -- a question which seems to be not very closely connected with our earlier computation of the sums $5 + 5 + 5$ and $3 + 3 + 3 + 3 + 3$.

By contrast, let us seek to verify the fact that $3 \cdot 5 = 5 \cdot 3$ using the second definition of multiplication. To compute $3 \cdot 5$ we first find two sets, say M and J, such that $n(M) = 3$ and $n(J) = 5$, and then we will have $3 \cdot 5 = n(M \times J)$. On the other hand, to compute $5 \cdot 3$ we must form the cartesian product $J \times M$ (since the number of elements in J is 5 and the number in M is 3), and then $5 \cdot 3 = n(J \times M)$. Will we find $n(M \times J) = n(J \times M)$, as claimed in this case by the commutative law of multiplication? Let us see.

A convenient set to use for M is $\{1, 2, 3\}$, since certainly $n(M) = 3$ in this case! Similarly, we may choose $J = \{1,2,3,4,5\}$ since, as we recall, the definition of multiplication does not require M and J to be disjoint. Then the elements of $M \times J$ will be all the ordered pairs (1,1), (1,2), (1,3), (1,4), (1,5), (2,1), (2,2), (2,3), (2,4), (2,5), (3,1), (3,2), (3,3), (3,4), and (3,5). On the other hand, the elements of $J \times M$ are the ordered pairs (1,1), (1,2), (1,3), (2,1), (2,2), (2,3), (3,1), (3,2), (3,3), (4,1), (4,2), (4,3), (5,1), (5,2), and (5,3). Notice that we can conclude that these two sets, $M \times J$ and $J \times M$, have the same number of elements -- without counting either set! The reason is that the ordered pairs which make up $J \times M$ are simply those of $M \times J$ "turned around" so that the first element of an ordered pair of $M \times J$ becomes the second element of a certain ordered pair of $J \times M$ and vice-versa. Thus $n(M \times J) = n(J \times M)$, and so $3 \cdot 5 = 5 \cdot 3$.

The phenomenon encountered in this example is quite general. If A and B are any sets whatever, we have $A \times B \neq B \times A$ -- unless A and B are the same set. However, in every case we have $n(A \times B) = n(B \times A)$, because the ordered pairs which make up $B \times A$ are simply those of $A \times B$ "turned around". (This can be seen by reviewing the definition of cartesian product, §1.) This fact leads at once to the commutative law of multiplication, in full generality. For if x, y are any two whole numbers, and if we choose sets A and B such that $n(A) = x$ and $n(B) = y$, then the second definition of multiplication tells us that $x \cdot y = n(A \times B)$ and that $y \cdot x = n(B \times A)$. Since, as we have just seen, $n(A \times B) = n(B \times A)$ by the "turned around" principle, we conclude by the logic of equality that $x \cdot y = y \cdot x$ -- as claimed in the commutative law.

The Associative Law for Multiplication. If x, y, z are any whole numbers then we have $(x \cdot y) \cdot z = x \cdot (y \cdot z)$.

The left side, $(x \cdot y) \cdot z$, represents the number obtained by first forming the product $x \cdot y$, and then multiplying this by z. The right side, $x \cdot (y \cdot z)$, represents the number obtained by first forming the product $y \cdot z$, and then multiplying x by it. (Compare item 1, §3.) The associative law asserts that these two processes lead to the same number -- no matter which whole numbers x,y,z we take. Let us try to see why this law is true, using our first definition of multiplication.

Consider, for example, the case where x is 2, y is 3, and z is 4, and let us look at the term on the right side of the equation in the associative law, $2 \cdot (3 \cdot 4)$. According to our first definition of multiplication $3 \cdot 4 = (4 + 4 + 4)$, and hence $2 \cdot (3 \cdot 4)$, which is $2 \cdot (4 + 4 + 4)$, must be $(4 + 4 + 4) + (4 + 4 + 4)$. In this last expression there are 3 occurrences of the numeral "4" within the first parentheses, and 3 within the second, so that altogether we are adding $2 \cdot 3$ occurrences of 4. Because of the associative law of *addition*, we can (as indicated in Chapter 3, §4) express this sum as $4 + 4 + 4 + 4 + 4 + 4$ without reference to any particular pattern of parenthesizing the five addition operations to be performed. Since we are adding $2 \cdot 3$ accurrences of 4, the number we get is $(2 \cdot 3) \cdot 4$, according to the definition of multiplication. Since we started with $2 \cdot (3 \cdot 4)$, we have the desired equality: $(2 \cdot 3) \cdot 4 = 2 \cdot (3 \cdot 4)$.

These observations are quite general. If x,y,z are any whole numbers -- neither x nor y being 0 -- then $y \cdot z$ is a sum of y occurrences of z. Hence $x \cdot (y \cdot z)$ is a sum of x expressions, each a sum of y occurrences of z. Thus the total number of occurrences of z which must be added to get $x \cdot (y \cdot z)$ is $y + y + \ldots + y$, where here we have x occurrences of y. In other words, (applying the definition of multiplication once more), we must add together $x \cdot y$ occurrences of z to get $x \cdot (y \cdot z)$. But adding $x \cdot y$ occurrences of z gives $(x \cdot y) \cdot z$, by definition of multiplication. Hence $(x \cdot y) \cdot z = x \cdot (y \cdot z)$, as claimed.

We must still consider the case where one of the numbers x or y is 0. Suppose, for example, that x is 0. Then $x \cdot (y \cdot z)$ is $0 \cdot (y \cdot z)$, and this is 0 by the special clause of the first definition of multiplication (which asserts that 0 multiplied by any whole number gives a product which is 0). Thus $x \cdot (y \cdot z) = 0$. But $x \cdot y = 0 \cdot y$ in the case where x is 0, and $0 \cdot y = 0$ by the special clause, so that $x \cdot y = 0$. It follows that $(x \cdot y) \cdot z$ is $0 \cdot z$. But $0 \cdot z = 0$, by another use of the special clause, so $(x \cdot y) \cdot z = 0$ by the logic of equality. Since we have shown both $x \cdot (y \cdot z) = 0$ and $(x \cdot y) \cdot z = 0$, we finally have $x \cdot (y \cdot z) = (x \cdot y) \cdot z$ for the case where x is 0, as claimed. The case where y is 0 can be handled similarly; we leave details to the reader.

The law of multiplicative identity: For any whole number x we have $x \cdot 1 = x$.

The truth of this law is not hard to see using either of our definitions of multiplication. It will be recalled that the ~~additive identity element is 0, since when 0 is added to any~~ given number the result is that same number. Our new law shows that the multiplicative identity element is 1, since when any given number is multiplied by 1 the result is that same number.

Of course the number 0 plays a special role in the theory of multiplication, too, as we see by the special clause of the first definition: For any whole number x we have $0 \cdot x = 0$ (and hence also $x \cdot 0 = 0$, since $x \cdot 0 = 0 \cdot x$ by the commutative law for multiplication.) There is no similar phenomenon in the theory of addition; that is, there is no whole number z such that for every number x we have $x + z = z$.

Because of this special role of 0 in the theory of multiplication, we can not have a cancellation law for multiplication of the same kind as we encountered in studying addition (Chapter 3, §4). In other words, it is possible to have whole numbers x, y, and z such that $x \cdot z = y \cdot z$, and yet $y \neq z$. For example, $2 \cdot 0 = 3 \cdot 0$ (since $2 \cdot 0 = 0$ and $3 \cdot 0 = 0$), but of course $2 \neq 3$. Indeed, if x and y are any two different whole numbers we will have $x \cdot 0 = y \cdot 0$ but $x \neq y$.

Although the cancellation law does not hold in full generality for multiplication, we have a modified form of it.

Limited Cancellation Law for Multiplication. If x, y, z are any whole numbers such that $x \cdot z = y \cdot z$, and if $z \neq 0$, then $x = y$.

How can we see that this general statement is true? Suppose that x, y, z are whole numbers such that $x \cdot z = y \cdot z$ and $z \neq 0$. Let us *assume*, temporarily, that $x \neq y$. Then one of the numbers x and y must be smaller than the other -- say $x < y$. From the equation $x \cdot z = y \cdot z$ we can obtain another by replacing the left side by a sum of x occurrences of z, and the right side by a sum of y occurrences of z. Now apply the *cancellation law for addition* x times, successively, to this equation. On the left side we will be left with 0, of course, but since $x < y$ we see that on the right side we will still have a sum of one or more occurrences of z. But since $z \neq 0$, by hypothesis, we *cannot* have 0 equal to a sum of one or more occurrences of z. We have thus arrived at a contradiction. This contradiction arises from our (temporary) *assumption* that $x \neq y$, and thus shows that after all we cannot have $x \neq y$. That is, we must have $x = y$ if we start with the hypotheses that $x \cdot z = y \cdot z$ and $z \neq 0$. This is the desired limited cancellation law.

Closure Law for Multiplication: *If x and y are any two whole numbers, then the product $x \cdot y$ is also a whole number.*

As we have indicated in the case of the corresponding law for addition, this fact about multiplication is part of what we mean by saying that multiplication is an operation on the set W of all whole numbers. Hence it follows immediately from either of our two definitions of multiplication.

We also express this fact by saying that the set *W is* *closed under multiplication*. Various subsets of W are also closed under multiplication, e.g., the set of all odd whole numbers, or the set of all positive whole numbers (i.e., all whole numbers other than 0).

We now turn to a very important law which connects multiplication with addition.

The distributive law of multiplication over addition. For any whole numbers x, y, z we have $x \cdot (y + z) = (x \cdot y) + (x \cdot z)$.

The left side, $x \cdot (y + z)$, is the number obtained by first adding y to z, and then multiplying x by this sum. The right side, $(x \cdot y) + (x \cdot z)$ is obtained by first forming the two products $x \cdot y$, and $x \cdot z$, and finally adding these two products together. The distributive law asserts that these two processes of computation always lead to the same result -- no matter what the numbers x, y, and z may be. How do we see that it is true?

If x is not 0, then the first definition of multiplication tells us that $x \cdot (y + z)$ is the sum $(y + z) + (y + z) + \ldots + (y + z)$, where we have added a total of x occurrences of the term $y + z$. Using the commutative and associative laws for *addition*, we can separate out the y's and the z's, getting $x \cdot (y + z) = (y + y + \ldots + y) + (z + z + \ldots + z)$, where on the right we are first adding x occurrences of y and then x occurrences of z. But applying our definition of multiplication again (twice), we see that the sum of x occurrences of y is $x \cdot y$ and the sum of x

occurrences of z is $x \cdot z$. Hence, we get $x \cdot (y + z) = (x \cdot y) + (x \cdot z)$, the desired d~~istribu~~tive law.

Let us close this section by citing laws which connect multiplication with the relation less than, $<$, and the successor operation S.

If x, y, z *are any whole numbers such that* $x < y$ *and* $z \neq 0$, *then* $x \cdot z < y \cdot z$.

For all whole numbers x, y *we have* $x \cdot S(y) = (x \cdot y) + x$.

§5 (Track B)

1. *Exercise*. Using the first (i.e., the repeated addition) definition of multiplication, show in detail that the associative law of multiplication, $(x \cdot y) \cdot z = x \cdot (y \cdot z)$, is also valid for the case $y = 0$. (Compare the case $z = 0$ treated in detail in §4.)

2. *Exercise*. Using the second (cartesian product) definition of multiplication, give a convincing argument for the validity of the multiplicative identity law.

3. *Exercise*. All sorts of operations on quite arbitrary sets may satisy some of the laws which we have seen hold for the operations $\cdot$ and $+$ on the set W. Assume that $\odot$ and $\boxplus$ are two-place operations on some set H. Express formally, using the letters "x", "y" and "z", the general statements that: (a) the set H is closed under $\odot$, (b) the operation $\odot$ is

associative, (c) the operation $\boxplus$ is commutative, and (d) $\odot$ is distributive over $\boxplus$.

4. Exercise. Let $J = \{1,3\}$ and let K be the set whose elements are all of the subsets of J. Thus K has exactly 4 elements. (a) List the 4 elements of K and label them a, b, c, and d.

Next, let us define the two-place operations $\oplus$ and $\otimes$ on the set K to be the operations of union and intersection, respectively; this makes sense, since the elements of K are sets. (b) Make a table of the 16 elementary facts about the operation $\oplus$ and another such table for the operation $\otimes$. Use the letters a, b, c, d for entries in the table. (c) Are either or both of the operations $\oplus$ and $\otimes$ commutative?
(d) Is either operation distributive over the other?
(e) Do there exist identity elements for the operations $\oplus$ and $\otimes$?
(f) Do either of the operations $\oplus$ and $\otimes$ satisfy a cancellation law?
(Justify your answers to (c), (d), (e) and (f) by referring to the tables prepared in answer to (b).)

3	4	12
2	1	2
6	4	24

5. Exercise. Consider the following array: The numbers in the boxes are first multiplied horizontally and vertically and the resulting numbers are placed in the margins at the right and at the bottom. These marginal numbers are then multiplied to give the same final product, entered in the box at the lower right, which for the

above array is 24. (a) Fill in the boxes to give another example and work out the products. Express the fact that the product of the numbers in the right margin has the same value as the product in the bottom margin by means of a general law, using four variables. Carefully explain why this general law is true by using the commutative and associative laws for multiplication. (b) Enlarge the array, horizontally with more rows and then vertically with more columns. Do the marginal products still "work"? Why?

6. Classroom Discussion. Discuss the cartesian product of any finite set with the empty set $\emptyset$ and relate this to the second definition of multiplication.

7. Individual project. In Chapter 3, §9, we considered an axiomatic approach to the theory of whole numbers. The axioms mentioned there were the Principle of Mathematical Induction (page 3.37), and Axioms 1 and 2 involving addition (page 3.39). Let us now enlarge this system by adding two axioms involving multiplication, as follows.

Axiom 3. For every x in W we have $x \cdot 0 = 0$.

Axiom 4. For every x, y in W we have

$$x \cdot S(y) = (x \cdot y) + x.$$

In the enlarged axiom system we will of course have the associative law for addition (proved on pp. 3.39 - 3.40), and the commutative law for addition (mentioned on p. 3.43). Using these, together with the new Axioms 3 and 4, prove:

(a) The distributive law for multiplication over addition,

(b) the associative law for multiplication, and

(c) the commutative law for multiplication.

8. Exercise. (a) Formulate a general statement expressing the proposition that the operation of addition is distributive over multiplication.

(b) Give an example to show that this law is false.

9. Exercise. Give an example of a subset of W which is not closed under multiplication.

10. Exercise. Using the first (i.e., the repeated addition) definition of multiplication, the distributive law for multiplication, and the fact that $1 + 1 + 1 = 3$, show that $3 \cdot 5 = 5 \cdot 3$.

11. Exercise. In discussing the distributive law for $\cdot$ over $+$, in §4, it was stated that for any whole numbers y and z,

$$(y + z) + (y + z) + \ldots + (y + z) = (y + y + \ldots + y) + (z + z + \ldots + z),$$

where in each case the 3 dots represent the same number of omitted terms. Consider the case of this law in which the term $(y + z)$ occurs only twice on the left, and prove this case using the commutative and associative laws of addition.

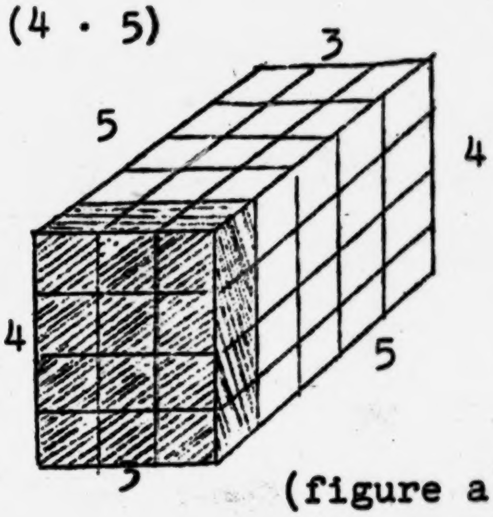
12. Exercise. At the end of §4 are given two laws connecting $\cdot$ with $<$ and with S. Find other true laws connecting $\cdot$ with $<$, $\leq$, and with S.

§6 (Track C)

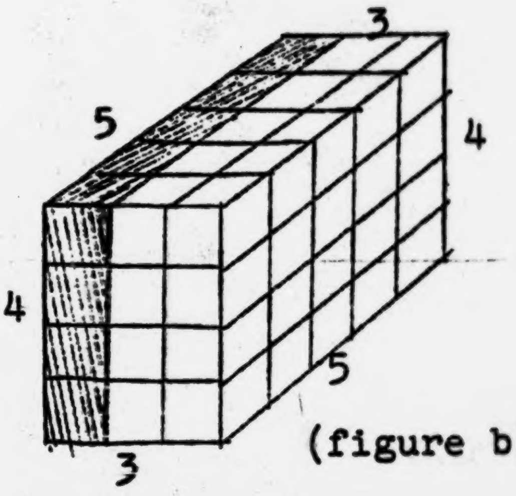
1. Learning Multiplication Facts. See exercise 5 of §5 for a practical method of teaching the elementary multiplication facts, which may also be used as a way of demonstrating the commutative and associative laws of multiplication. Such tables can be introduced at first as a "game". Later they can be used to "check" computation of products of 4 whole numbers. Note that the associative law for multiplication is itself a method for "checking" products of 3 whole numbers.

2. The Associative law of Multiplication can be visually grasped by making (or ordering from the school district, if lucky) a rectangular box made up of unit cubes, or blocks, which can be fitted together.

For example, to show $(3 \cdot 4) \cdot 5 = 3 \cdot (4 \cdot 5)$ use a model as in figure (a) where we have $3 \cdot 4$ blocks in each vertical slab and there are 5 such vertical slabs. This illustrates the left side of the above equation. For the right side the box can be looked at as in figure (b) where there are $4 \cdot 5$ boxes in each vertical slice and 3 such slices. Of course, no matter how we slice it, we have the same number of little blocks in the big box.



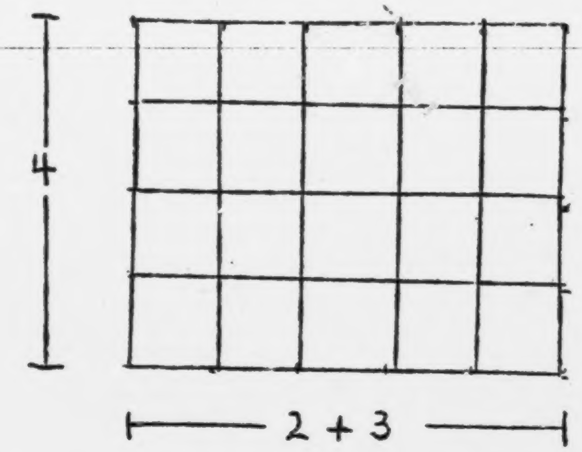
(figure a)



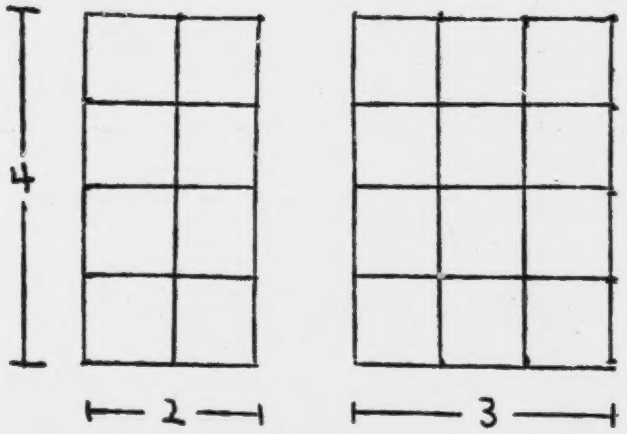
(figure b)

3. The distributive law for multiplication over addition. This law can also be displayed visually in a way which makes it intuitively simple to grasp. In contrast to the associative law discussed in item 2 above, only two-dimensional squares are needed instead of three-dimensional blocks.

For example, to illustrate that $4 \cdot (2 + 3) = (4 \cdot 2) + (4 \cdot 3)$, we consider the 4×5 array of squares



which contains $4 \cdot (2 + 3)$ squares. By cutting down a vertical line and separating the two pieces we get



The split array has $(4 \cdot 2) + (4 \cdot 3)$ squares in it. Since the number of squares was not changed by cutting and separating the original array, we see that $4 \cdot (2 + 3)$ is the same as $(4 \cdot 2) + (4 \cdot 3)$. Notice that it is not necessary to evaluate the total number of squares as 20 in order to come to the conclusion that $4 \cdot (2 + 3) = (4 \cdot 2) + (4 \cdot 3)$!

4. The use of laws in computation. The laws of addition and multiplication should be learned by elementary school students---preferably by discovery rather than by being told what the laws are. Motivation to learn the laws, a reinforcement of the learning process, should be brought about by indicating how the laws may be used to simplify computations. For example, if one is asked to compute $(13 \cdot 5) \cdot 4$ and goes at it in a straightforward way, one first gets $13 \cdot 5 = 65$ and then $65 \cdot 4 = 260$; neither of these multiplications can be done "in the head" by beginning students. However, if one converts the given $(13 \cdot 5) \cdot 4$ into $13 \cdot (5 \cdot 4)$ by the associative law for multiplication, then the products which must be calculated are $5 \cdot 4 = 20$ and $13 \cdot 20 = 260$, both of which are much easier. Make up other examples of this kind, in which a combination of commutative and associative laws can be used to simplify computations.

Especially useful is the distributive law. For instance, $8 \cdot 13$ can be written as $8 \cdot (10 + 3)$ which, by the distributive law is the same as $(8 \cdot 10) + (8 \cdot 3)$ or $80 + 24$, which is 104. Of course, it is just such a use of the distributive law which underlies the algorithm for multiplying with 2-place numerals, as we shall see in Chapter 5. Notice, however, that the distributive law can be used in many different ways to evaluate a given product such as $8 \cdot 13$. For example,

$$8 \cdot 13 = 8 \cdot (8 + 5) = (8 \cdot 8) + (8 \cdot 5) = 64 + 40 = 104.$$

§7 (Track A)

In this section we shall investigate some of the geometric pictures connected with the operation of multiplication.

Consider first the commutative law for multiplication. In verifying the special case of this law, $3 \cdot 5 = 5 \cdot 3$, we considered the sets $M = \{1,2,3\}$ and $J = \{1,2,3,4,5\}$, we formed the two cartesian product sets $M \times J$ and $J \times M$, and finally we saw that these two product sets, while not the same, have the same number of elements. This last fact can be seen pictorially by constructing a pair of coordinate axes in a geometric plane, and considering the lattice points represented by the ordered pairs of $M \times J$ and of $J \times M$.

In fact, since the elements of $M \times J$ are all those ordered pairs (x,y) obtained by taking x to be any of the numbers $1,2,3$ and y to be any of $1,2,3,4,5$, we see that the points of the plane represented by these elements form a rectangular array consisting of 3 columns and 5 rows. (See Figure 1.)

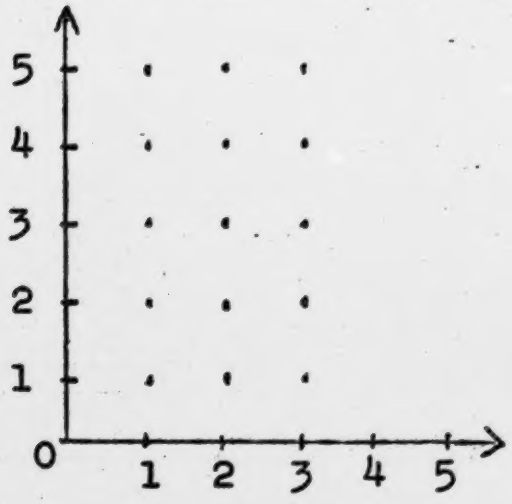


Figure 1.

On the other hand, when we mark the points corresponding to the elements of $J \times M$, as in Figure 2, we find a rectangular array consisting of 5 columns and 3 rows.

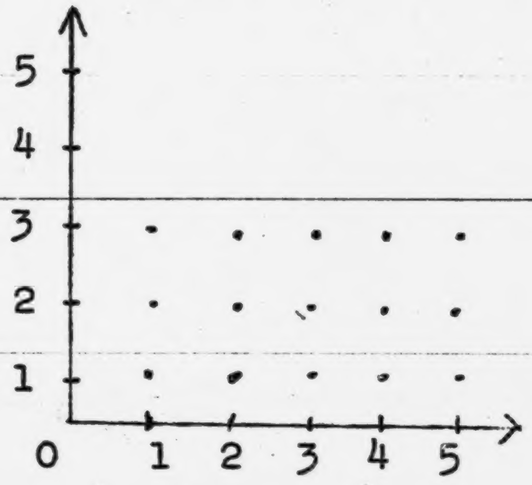


Figure 2.

It is geometrically evident that the arrays in Figures 1 and 2 have the same number of points, for either one of these arrays can be obtained from the other by a process of rotating and sliding.

It is useful to note that the elements of any cartesian product of sets, $A \times B$, can be represented by a rectangular array, whether or not the elements of the sets A and B are numbers. For example, if A = {George Washington, Abraham Lincoln}, and if B = {New York, Los Angeles, Kansas City, Chicago}, and if we wish to represent $A \times B$, we select 2 points on a horizontal axis which we label "George Washington" and "Abraham Lincoln", we select 4 points on a vertical axis which we label "New York, "Los Angeles", "Kansas City", and "Chicago", and then each of the 8 lattice points determined by these 6 selected points will correspond to one of the ordered pairs making up the cartesian product $A \times B$. These lattice points form a rectangular array of 2 columns and 4 rows. In Figure 3

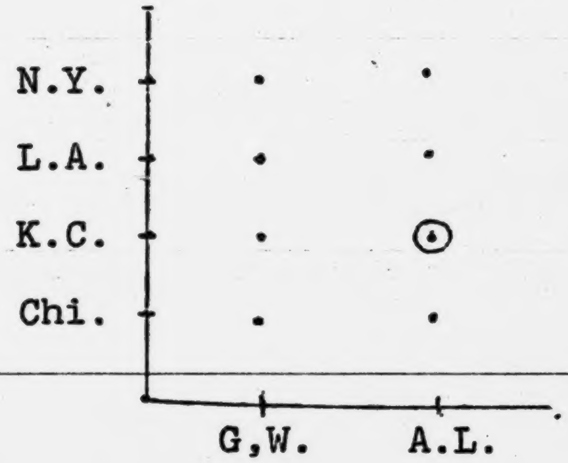


Figure 3.

we have circled the point (Abraham Lincoln, Kansas City). Using the same sets A and B, we would picture the cartesian product $B \times A$ by labelling 4 selected points on the horizontal axis with the elements of B, and 2 selected points on the vertical axis with the elements of A. The picture of $B \times A$ then consists of a rectangular array having 4 columns and 2 rows.

This way of picturing cartesian products makes clear that for any finite sets C and D we have $n(C \times D) = n(D \times C)$, a principle which underlies the commutative law of multiplication. But these pictures of cartesian products can also help us to see the truth of another proposition about sets, which leads to an alternative method of understanding the distributive law of multiplication over addition.

Proposition. If A,B,C are any sets such that $B \cap C = \emptyset$ (i.e., such that B and C are disjoint), then also $A \times B$ and $A \times C$ are disjoint, and we have

$$A \times (B \cup C) = (A \times B) \cup (A \times C).$$

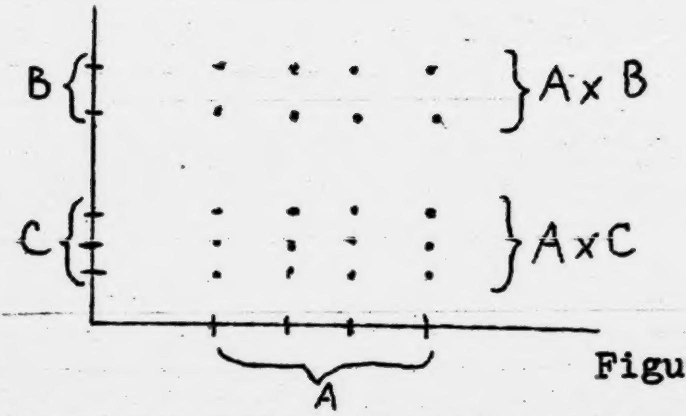


Figure 4.

Let A, B, and C be given sets, B and C disjoint. Following the pattern outlined above, we picture the elements of A as points on a horizontal axis, and the elements of B and of C as points on a vertical axis; since B and C are disjoint, we may place all the points corresponding to elements of B above the points corresponding to elements of C. Now when we mark the lattice points determined by the selected points on the axes, we see that the full rectangular array, which represents the elements of $A \times (B \cup C)$, breaks naturally into two disjoint subsets -- the lattice points representing $A \times B$ above, and the lattice points representing $A \times C$ below. This illustrates the two parts of the conclusion of our Proposition: $A \times B$ and $A \times C$ are disjoint, and $A \times (B \cup C) = (A \times B) \cup (A \times C)$.

Now using the Proposition we have just illustrated, we may obtain a proof of the distributive law of multiplication over addition as follows:

1. Let x, y, z be any whole numbers.
2. Choose disjoint set B and C so that $n(B) = y$ and $n(C) = z$.
3. $y + z = n(B \cup C)$; by line 2 and the first definition of addition (Chapter 3, §1).

4. Choose any set A such that $n(A) = x$.
5. $x \cdot (y + z) = n(A \times (B \cup C))$; by lines 3, 4 and the second definition of multiplication (§1).
6. $A \times (B \cup C) = (A \times B) \cup (A \times C)$; by Proposition above.
7. $x \cdot (y + z) = n((A \times B) \cup (A \times C))$; by lines 5, 6 and logic of equality.
8. $(A \times B)$ and $(A \times C)$ are disjoint; by Proposition above, since B and C are disjoint by line 2.
9. $n(A \times B) = x \cdot y$ and $n(A \times C) = x \cdot z$; by lines 2, 4 and second definition of multiplication.
10. $n((A \times B) \cup (A \times C)) = (x \cdot y) + (x \cdot z)$; by lines 8, 9 and first definition of addition.
11. $x \cdot (y + z) = (x \cdot y) + (x \cdot z)$; by lines 7, 10 and logic of equality.

Since x, y, z are any whole numbers (line 1), we see that line 11 establishes the distributive law.

Let us now look at the pictures of graphs of equations involving the operation of multiplication. If we first look at the graphs of the equations $y = 1 \cdot x$, $y = 2 \cdot x$, $y = 3 \cdot x$, we see that each graph lies along a straight line, and that the lines corresponding to the three equations get successively steeper. (See Figure 5.) On the other hand, the graphs of the equations $x = 1 \cdot y$, $x = 2 \cdot y$, $x = 3 \cdot y$ lie along lines which get successively less steep. (See Figure 6.)

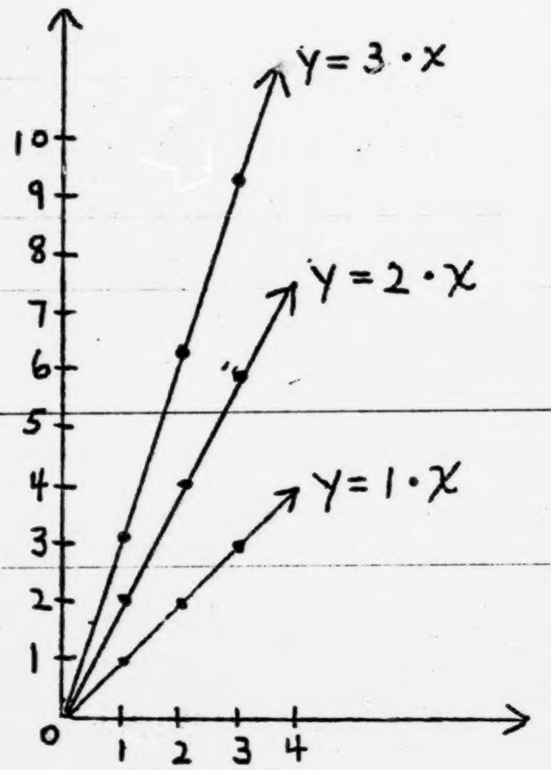


Figure 5.

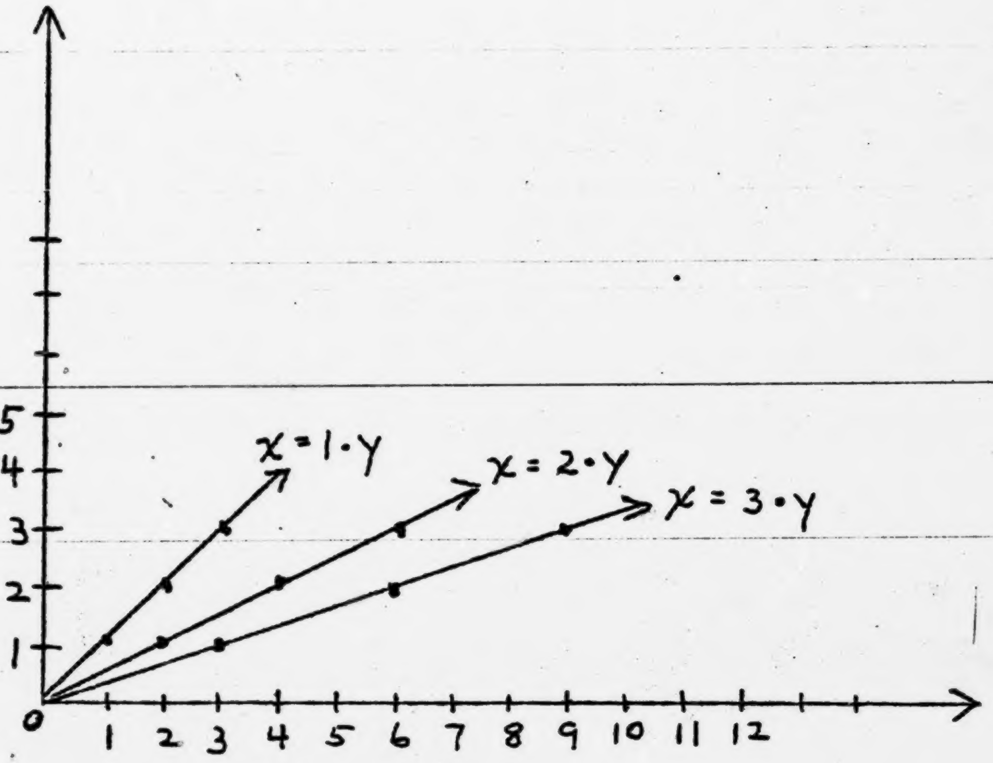


Figure 6.

Finally, the graphs of each of the equations $x \cdot y = 1$, $x \cdot y = 2$, $x \cdot y = 4$, $x \cdot y = 6$ have only a finite number of lattice points, and those that contain more than 2 points do not lie on a single straight line. (See Figure 7.)

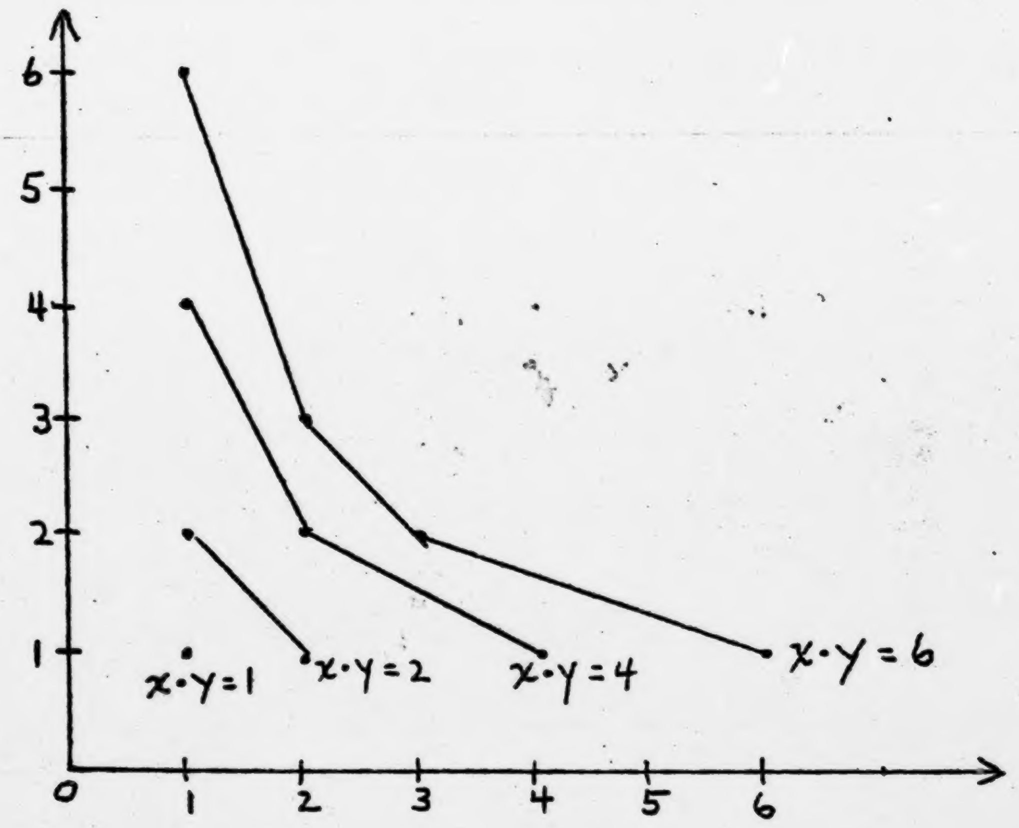


Figure 7.

Another type of geometric representation for equations, as we have seen in Chapter 3, §6, involves motions of the number line. For example, given the multiplicative equation $y = 2 \cdot x$, we consider a whole number line and from each point with coordinate x on this line we draw an arrow pointing to the point $2 \cdot x$.

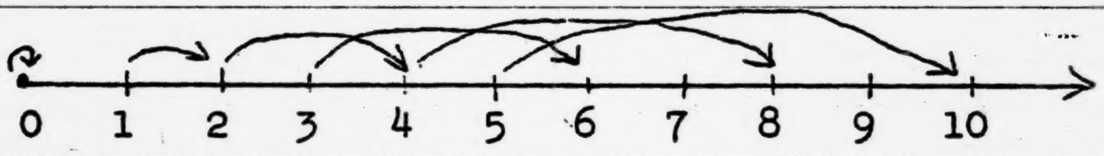


Figure 8.

The pattern of these arrows suggests a motion of this line -- one in which the endpoint, 0, does not move at all, and in which the points further from 0 move further during the motion. (Figure 8.) If we indicate the position of the points before the motion on one number line, and the position of the same points at the end of the motion on another line right below the first one, the picture we get (Figure 9) suggests that the motion is a uniform stretching of the line. In this stretching, each

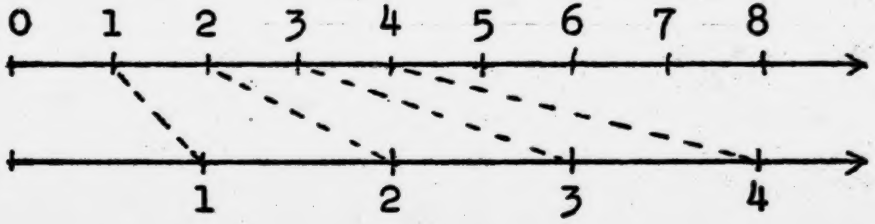


Figure 9.

point moves to the right a distance equal to its original distance from the 0 point. Similarly, the equation $y = 3 \cdot x$ can be pictured as a stretching motion of the number line in which each point moves twice as far to the right as its original distance from the 0 point.

Figure 9 has another interpretation. Instead of the two pictured lines being thought of as the before-and-after pictures of a single line in motion, we can simply consider what we actually see -- two half lines side by side. On each we have laid off a number line, but the unit distance on the upper line is such that twice its length can be fitted into the unit distance on the lower line. When the two number lines are laid next to one another in this fashion, to each number x on the lower line we can read off the value $2 \cdot x$ on the upper line immediately above. Similarly, by starting with a number line and placing below it another number line whose unit distance corresponds to the number 3 on the first line, we can look up any number x on the lower line and find $3 \cdot x$ directly above. This connection with change-of-units is one of the important areas of application of the multiplication operation.

To conclude this section, we wish to indicate that a multiplicative equation such as $y = 2 \cdot x$ can be represented by a certain rigid motion of a number line, as well as by a stretching motion. What we have in mind is a rotation. Indeed, if we start with a number line in horizontal position and rotate it counter-clockwise, at a certain position the point 2 on the line will lie directly above the original location of the point 1. (Figure 10.) If we stop the motion at that position, then for every whole number x, the point $2 \cdot x$ will lie directly above the original location of the point x.

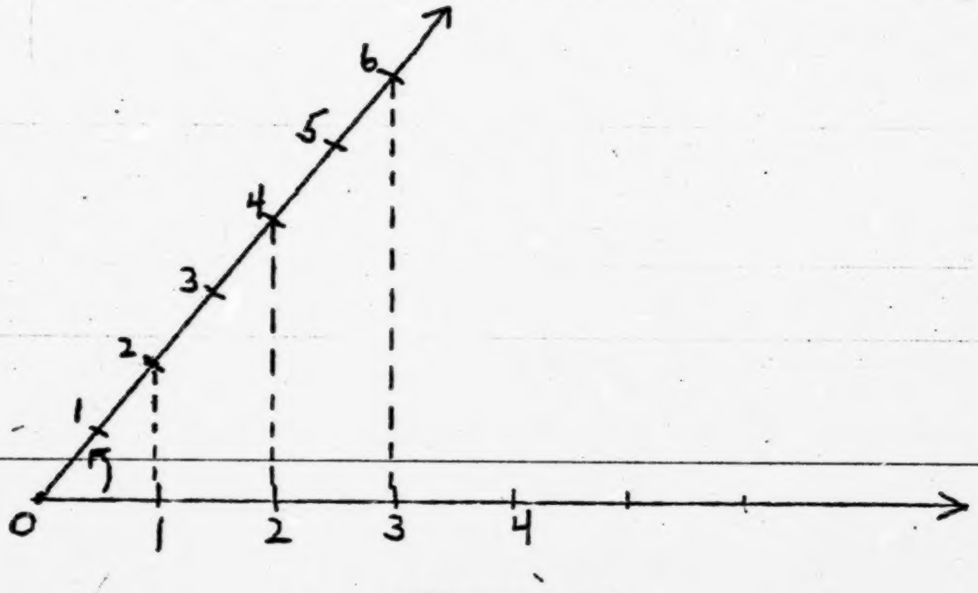


Figure 10.

The reason for this is the geometric fact of proportionality in similar triangles. For example, in Figure 10 the triangle whose vertices are (0, upper 2, lower 1) and the triangle (0, upper 6, lower 3) are similar, because two sides of the small triangle are on the same lines as the corresponding sides of the large triangle, while the third sides of these two triangles are parallel (both being vertical). The geometric theory of similar triangles then tells us that the ratio of the lengths of the bottom side to the top side of the small triangle, must be the same as the ratio of the lengths of the bottom side to the top side of the large triangle.

§8 (Track B)

1. Exercise. Suppose that C = [(California, Sacramento), (Oregon, Salem), (Washington, Sacramento), (Oregon, Sacramento), (California, Salem), (Washington, Salem)]. List the elements of sets A and B such that $A \times B = C$. Illustrate the elements of C by means of a rectangular array.

2. Let H be the set of those ordered pairs (of whole numbers) which are represented by the lattice points in figure (a) below. (i) Find sets A,B,C,D,E,F of whole numbers, such that $H = (A \times D) \cup (B \times E) \cup (C \times F)$. (ii) Do we also have $H = (A \cup B \cup C) \times (D \cup E \cup F)$? Justify your answer. (iii) Are the sets A and B disjoint? If so, can they be replaced by other sets A' and B' which are not disjoint in satisfying (i)?

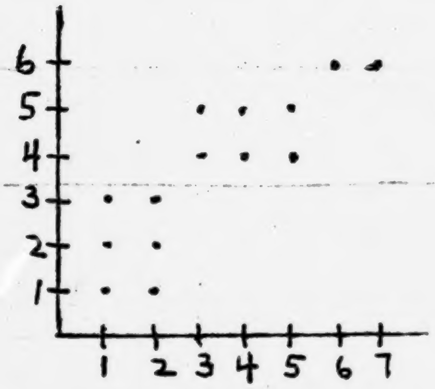


figure (a)

3. Exercise. For each of the following equations, state whether its graph consists of a finite or infinite number of lattice points, and whether it lies on a single straight line or not:

(i) $x \cdot y = 3$

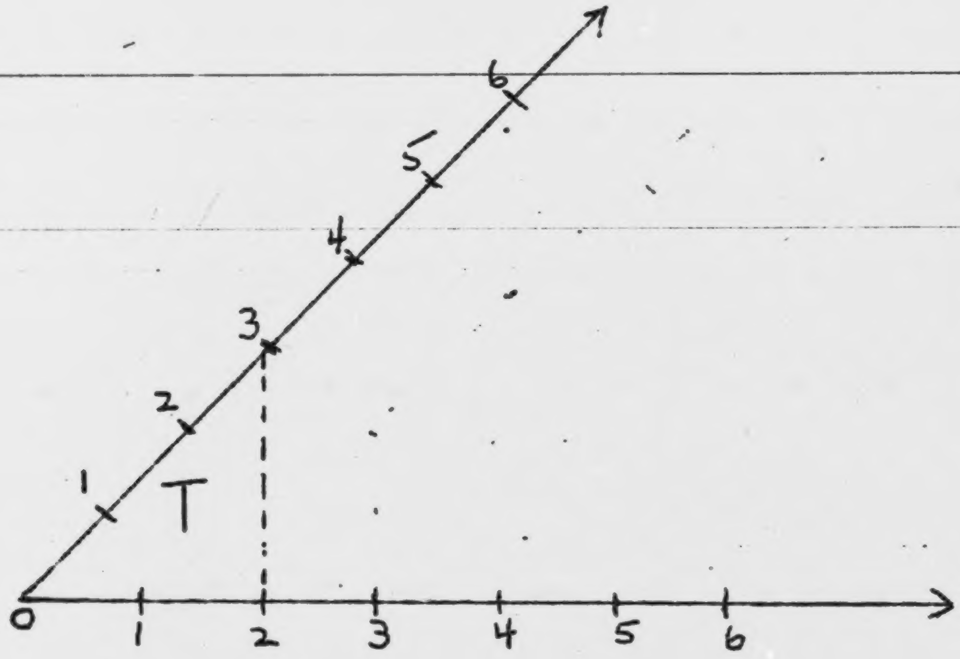
(ii) $x = 5 \cdot y$

(iii) $y \cdot x = 16$

(iv) $4 \cdot x = y$.

4. Exercise. What equation can be pictured geometrically as a uniform stretching motion of a number line on which each labeled point moves three times as far to the right as its original distance from the zero point?

5. Exercise. Consider the following figure on which a triangle T with vertices (0, upper 3, lower 2) is described. (a) Give the vertices of two more triangles, U and V, each of which is similar to the given one.

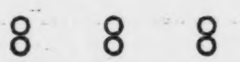


(b) Are these two triangles you found similar to each other? (c) Write an equation involving variables "x" and "y", and two uses of the multiplication sign, such that when a numerical value x is given the corresponding number y can be read off by means of the above diagram. Explain how we do this "reading off".

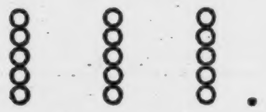
6. Discussion. As in Exercise 4, certain equations lead to "stretching" of a line. Discuss stretching and shrinking generally. Are such motions of a line "rigid motions"? What is a rigid motion?

§9 (Track C)

1. The Distributive law of Multiplication. As indicated in §3, item 2, we can proceed, without mention of cartesian products, to consider rectangular arrays of physical sets in the classroom. These are very helpful for teaching multiplication and related laws. For example, to show $3 \cdot (2 + 5) = (3 \cdot 2) + (3 \cdot 5)$, obtain a bag of marbles or plastic chips. Then the product $3 \cdot 2$ can be pictured by arranging the chips as follows:

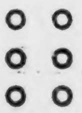


To get the product $3 \cdot 5$ arrange them:

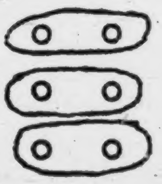


Hence $(3 \cdot 2) + (3 \cdot 5)$ is illustrated by the entire array. But by moving the top batch down an inch or so, we see that we have just $3 \cdot (2 + 5)$ chips.

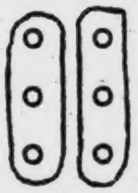
2. The Commutative law of multiplication is even easier to demonstrate using physical objects, since to show, for example, that $3 \cdot 2 = 2 \cdot 3$, arrange the chips in rectangular array, such as



Then circle these chips with a crayon to show 3 sets of 2 or 2 sets of 3.



or



3. Change of units. Measuring physical objects by using a number line can be a stimulating classroom activity which helps to motivate the concepts involved in changes of units.

Consider a length of string which when stretched out and placed on a number line falls between two points on the line. How can we describe its length? One way would be to change the basic length of the unit on the number line, lengthening it or shortening it the proper amount so that the string, when placed in the new number line, would fall on a labeled point. For more advanced students consider the problem of devising a new number line (actually the basic unit) which is capable of measuring two different lengths of string, neither of which falls on a labeled point of the original number line.

MATHEMATICS 15

Spring Quarter 1968

Lectures by
Professor Leon Henkin

Notes by
Arthur Kessner

Other relations whose definitions should be clear are:

(i) less than or equal to, written $\leq$,

(ii) greater than or equal to, written $\geq$..

An Experiment in Elementary School Mathematics Instruction.

Going on in some elementary schools of Berkeley and certain other East Bay cities is an experiment in mathematics education which deserves the attention of anyone who plans to teach. This project, called by the acronym SEED, Special Elementary Education for the Disadvantaged, is operating on grade levels one through six in special schools, those which have a preponderance of students whose upbringing has been called disadvantaged (for a variety of reasons).

The project departs from traditional mathematics instruction along three main paths:

(1) The subject matter presented is the kind usually considered "advanced mathematics", for it revolves about algebra and abstract geometry. This is based on the remarkable realization that students in the elementary grades can comprehend the kind of mathematics ordinarily reserved for high school and college students.

(2) The teacher is not the usual elementary school teacher but rather a specialist in mathematics, one who has training at least at the level of B.A. in mathematics.

(3) The method these specialists are using is called the Discovery Method, whereby the student is directly told as little as possible, but instead he is led to make discoveries for himself.

The regular teacher remains in the room while a specialist takes over the class. This takes care of discipline problems and satisfies a state law concerning credentialed employees.

More importantly, however, this regular teacher is a witness to some remarkable transformations in her students. Positive motivation, so critical a concept in teaching and one so often lacking in average teaching becomes a reality when students are permitted and encouraged to creatively take part in the subject matter. The children become fascinated with the process of discovering and communicating mathematics. In more common language they are turned-on.

As an extension of the SEED program, in the ninth grade of Roosevelt Junior High in Oakland a math specialist has been working with the students, having them teach third graders advanced mathematics using the discovery method. This program has altered their lives to a considerable degree, and it is these ninth graders who will visit your sections at their next time of meeting and present a few facets of themselves and the SEED program.

Lecture 12, April 26, 1968

Remark: From the teaching demonstrations we have all witnessed, the possibilities for creative teaching are seen to be vast. We must rededicate ourselves to revolutionalizing outdated methods of teaching.

Recall that in our study of the Whole number system (W,+,·,exp), we began by describing a scheme for naming the elements of W. This was necessary since W is infinite, whereas the number of symbols employed to name the elements of W is finite. In our previous scheme we took ten basic symbols for the names of the first ten numbers. These were: 0, 1, 2, ..., 9. To name those numbers that followed nine we used combinations of these first ten names--that is, 10, 11, ..., 20, 21, ..., 30, ..., 90, ..., 99, 100, ... There is, however, no mathematical reason why we give different names to the first ten numbers and from these develop a scheme. In fact, there are other possible ways, each of which results in a different numeration scheme.

We'll now demonstrate one alternative numeration scheme, choosing 3 as its base--that is, we'll use three different basic symbols 0, 1, 2, and name all of W using combinations of these three. Thus, the whole numbers in their natural order would be named as follows:

0, 1, 2, 10, 11, 12, 20, 21, 22, 100, 101, 102, 110, 111, 112, 120, 121, 122, 200, 201, ...

Intuitively, you should be able to see that this is the same kind of scheme we have previously used, except that we have but 0, 1, 2 to work with. However, if we want to talk about both of these schemes, we have to have a method for distinguishing the two. One such method, the one we will adopt, makes the notational convention that numbers in the base 3 numeration scheme are to be put in parentheses with a subscript denoting the base.
E.g. $0, 1, 2, (10)_3, (11)_3, (12)_3, (20)_3, (21)_3, (22)_3, \ldots$

Now, considering only what we have called the counting numbers (that is, leaving out 0) we see that $(10)_3$ is the third number in the scheme of whole numbers and so $(10)_3 = 3$.

Similarly $(102)_3 = 11$

We'll now discuss how to translate from one numeration scheme to another without having to go through the numbers in their natural order. This translation can be accomplished by using the polynomial representation of numbers, which we have discussed before.

Recall that in the base 10 scheme, a number say 127, could be expressed in powers of the base 10,

i.e. $127 = (1 \cdot 10^2) + (2 \cdot 10^1) + (7 \cdot 10^0)$

There is a similar representation that can be made for numbers described in different bases.

For example,

$$\begin{aligned} (102)_3 &= (1 \cdot 3^2) + (0 \cdot 3^1) + (2 \cdot 3^0) \\ &= (1 \cdot 9) + (0 \cdot 3) + (2 \cdot 1) \\ &= 9 + 0 + 2 \\ &= 11 \end{aligned}$$

Hence $(102)_3 = 11$ which we knew by the simple process of counting the numerals in the base 3 scheme until we came to the eleventh, beginning as we mentioned before at 1.

With this process of representation in mind, we can find the way to express 127 in the base 3 number system.

First we represent 127 as a polynomial in powers of 3, using the following powers of 3 to help us.

$3^0 = 1$
$3^1 = 3$
$3^2 = 9$
$3^3 = 27$
$3^4 = 81$
$3^5 = 243$

Notice that 127 is smaller than 243 and so cannot be a multiple of 3^5. So we try 3^4

i.e. $127 = 1 \cdot 3^4 + \text{(something else)}$

The (something else) is found by first finding out how much is left in 127 after taking out $1 \cdot 3^4$; that is, after subtracting 81 from 127. This leaves 46.

Now out of 46, we can get $1 \cdot 3^3$ + (something else)

i.e. $127 = 1 \cdot 3^4 + 1 \cdot 3^3 + \text{(something else)}$

Doing what we did above, we finally arrive at the following:

$$127 = (1 \cdot 3^4) + (1 \cdot 3^3) + (2 \cdot 3^2) + (0 \cdot 3^1) + (1 \cdot 3^0)$$

hence $127 = (11201)_3$

Now, say we pick 2110 in base 3 (that is, $(2110)_3$) and want to express it in base 10. This we can do as follows:

$$\begin{aligned}(2110)_3 &= (2 \cdot 3^3) + (1 \cdot 3^2) + (1 \cdot 3^1) + (0 \cdot 3^0) \\ &= (2 \cdot 27) + (1 \cdot 9) + (1 \cdot 3) + (0 \cdot 1) \\ &= (54) + (9) + (3) + (0) \\ &= 66\end{aligned}$$

Thus $(2110)_3 = 66$

In this way you see it's possible to pass back and forth between the two numeration schemes. Either one, of course, is satisfactory in itself, and although you are more used to the base 10 scheme, our counting process, our algorithms and our general laws all hold in the base 3 system.

For example; say we wanted to count the elements in the set A = {R.Kennedy, R.Nixon, E.McCarthy, N.Rockefeller, D.Gregory, North Star}
following our procedure for counting, we take the elements of A and line them up with the counting numbers, this time using the base 3 system.

So,

D.G.,	N.R.,	N.S.,	E.M.,	R.N.,	R.K
1	2	$(10)_3$	$(11)_3$	$(12)_3$	$(20)_3$

Thus $n(A) = (20)_3$

If we had used our base 10 scheme, we would have found $n(A) = 6$, and, of course, this checks since $6 = (20)_3$.

Now, let's indicate how we would use the addition algorithm using the base 3 scheme. First, recall that this algorithm depended upon our use of the elementary addition table,

+	0,...,9
0	
⋮	
9	

In the base 3 system, elementary addition facts are even simpler.

+	0	1	2
0	0	1	2
1	1	2	$(10)_3$
2	2	$(10)_3$	$(11)_3$

This table thus becomes hypothesis (1) for the addition algorithm.

Hypothesis (2), our general laws among which are the associative and commutative laws, are exactly the same as before since these laws concern themselves with the whole numbers themselves and not the names of the numbers--that is, not the numeration scheme.

Hypothesis (3), the last of those underlining the addition algorithm, was the polynomial representation of a number in powers of the base.

Thus, to add $(201)_3$ and $(122)_3$ we proceed similarly to the way we did with numbers in the base 10 scheme. We "carry" when the numbers added are greater than 2, using the elementary addition table for the base 3 system.

$$\begin{array}{r} (201)_3 \\ \underline{(122)_3} \\ (1100)_3 \end{array}$$

We now check this result by using the translation process which we described above.

$$\begin{aligned} (201)_3 &= (2\cdot 3^2) + (0\cdot 3^1) + (1\cdot 3^0) \\ &= 18 + 0 + 1 \\ &= 19 \\ (122)_3 &= (1\cdot 3^2) + (2\cdot 3^1) + (2\cdot 3^0) \\ &= 9 + 6 + 2 \\ &= 17 \\ \text{Finally, } (1100)_3 &= (1\cdot 3^3) + (1\cdot 3^2) + (0\cdot 3^1) + (0\cdot 3^0) \\ &= 27 + 9 + 0 + 0 \\ &= 36 \end{aligned}$$

and since it is true that 19 + 17 is indeed equal to 36, our check is complete.

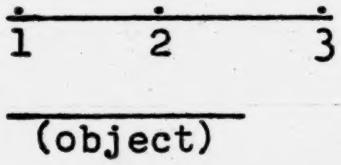
Before leaving numeration schemes, we might mention that certain schemes other than the base 10 have found applications outside of mathematics. In particular, the base 2 scheme, a most economical system because but two different symbols, 0, 1, are used, is important since it is this scheme which is used in almost all electronic computers.

In order to see why number systems other than (W, +, ·, exp) are desirable for study, let's examine a few of the shortcomings of the Whole Number system. From a practical standpoint, simple measurement of physical objects cannot be handled adequately in (W, +, ·, exp). Recall that we discussed representing the whole numbers on a line:

$$\begin{array}{ccccc} \cdot & \cdot & \cdot & \cdot & \cdots \\ 1 & 2 & 3 & 4 & \ldots \end{array}$$

But if we want to measure an object, its beginning or end may not coincide with a whole number.

e.g.



In addition, mathematicians themselves are not esthetically satisfied with the whole number system. For example, consider the process called subtraction, usually denoted by - placed between certain whole numbers. Mathematicians do not as a rule call - an operation on W because it cannot be applied to any pair of whole numbers x, y to get a new whole number x - y. In other words, subtraction does not satisfy a closure law in W. How could we define subtraction more satisfactorily? First, notice that it is a function which operates on certain ordered pairs (x, y) of whole numbers.

Let's define S to be the set of all those ordered pairs (x, y) of whole numbers such that $x \geq y$ i.e. such that the first member of the ordered pair is greater than or equal to the second member.

Using our definition of the relation $\geq$ and the basic laws we have studied, we could now prove the

Theorem: If (x, y) is any element of S then there is one and only one whole number z such that $x = y + z$.

As a result of this theorem, we could introduce subtraction into a subset of W, the set S, as follows.

Definition of Subtraction: For any ordered pair (x, y) in S, we define x - y as the unique whole number z such that $x = y + z$.

Lecture 13, April 29, 1968.

Recall from last lecture, we were discussing a few disadvantages of introducing subtraction into the whole number system $(W,+,\cdot,\exp)$; for example, when we try to extend the general laws which apply to $+,\cdot,\exp$ in an analogous way to subtraction, all kinds of "messy" exceptions are necessary. The following will further explain what we mean by "messy exceptions".

First we defined S as the set of all ordered pairs (x,y) of whole numbers such that $x \geq y$. Then we defined subtraction, $-$, as a function which acts on any pair (x,y) in the set S. The result of this action is a whole number z; namely, the unique whole number z such that $x = y + z$.

If, however, x,y are whole numbers such that the ordered pair (x,y) is *not* in S (i.e. if $x < y$), then there is *no* whole number z such that $x = y + z$, and hence subtraction is not defined for such an (x,y). In other words, $x - y$ is *meaningless* if $x < y$.

With this definition of subtraction in mind, lets now look at a few of the more familiar aspects of it.

There are *Particular Subtraction facts*, such as:

$12 - 4 = 8$
$3 - 0 = 3$
$5 - 4 = 1$
$4 - 5$: meaningless

Also, there are *General Statements* concerning subtraction:

e.g. For any whole number x in W, $x - x = 0$
For any whole number x in W, $x - 0 = x$

Lets check the possibility of a commutative law for subtraction in W. It would read:

for any x,y in W $x - y = y - x$

Clearly, this would only be true in case $x = y$, because if $x < y$, then by definition the left side of the above equation is *meaningless* whereas if $x > y$ the right side is *meaningless*.

Thus, there is no commutative law for subtraction in W.

What about an Associative law?

It would read: For any x, y, z in W $x - (y - z) = (x - y) - z$

Is this correct? No, it is false and we can show this by a particular example.

Take $x = 4$, $y = 2$, $z = 1$

Then the above equation reads $4 - (2 - 1) = (4 - 2) - 1$

That is, $4 - 1 = 2 - 1$ clearly false.

How can we change this law to make it a true one. One possible way is to make it read:

For any x, y, z in W $x - (y - z) = (x - y) + z$

Now, it is true that there are no numbers in W which make this equation false, but if $x = 0$ and $y > 0$ then the right hand side is meaningless; we have to search further.

We could again change it to read:

For any x, y, z in W if $x \geq y$ and $y \geq z$ then
$x - (y - z) = (x - y) + z$

Here finally, we have a true general statement, but not an associative law for subtraction. Notice, that to make it true, we were forced to impose fairly complicated conditions on x, y, z. This is what we meant when we described the general laws concerning subtraction as unesthetic.

Lets look at a distributive law for multiplication over subtraction, with the necessary conditions to make it a true statement.

For any x, y, z in W, if $y \geq z$,
then $x \cdot (y - z) = (x \cdot y) - (x \cdot z)$

It is precisely these messy conditions we will eradicate when we pass from the whole numbers to the integers.

> Remark: Usually, in the elementary schools, the positive rational numbers (fractions) are studied prior to the introduction of the negative numbers, but since there is nothing absolute about the order

in which they are presented, here we will go to the negative numbers first.

From our study of the whole numbers, we now know that: if a, b are any whole numbers such that $a \geq b$, then there is a whole number x such that $a = b + x$. But if a, b are whole numbers with $a < b$, then there is no whole number x such that $a = b + x$. In particular, there is no whole number x such that $0 = 1 + x$.

We are now going to extend our system of whole numbers to a new one with more numbers in it in which there will be a number (not a whole number) which when added to 1 will give 0.

This means:

(i) We must find a set of numbers, call it J, containing W as a subset with at least one new number in it.

(ii) Since the operation addition has only been defined in W, in order to add numbers in this new set J, which includes W as a subset, we must extend our previous definition of addition to include all the numbers in J. This means finding an operation, lets denote it by $+_J$, which can act on any numbers x, y in J with the result of this action being another number $x +_J y$ in J. Furthermore, we require that whenever x, y are in W, then $x +_J y = x + y$. That is, when this new operation $+_J$ is restricted to the numbers in W, it gives the same results as $+$. This is what is meant when we say $+_J$ is an extension of $+$.

(iii) Similarly for multiplication in the new set J, which we'll denote by $\cdot_J$.

Now, that we've described our desired goals, how can we accomplish them? There are, naturally, several equally valid ways. Moreover, as with any mathematical theory, there are axiomatic ways as well as the definitional approach. We have mentioned this earlier with reference to Euclidean Geometry and the axiomatic approach that was used by G. Peano.

To begin, we'll give a general idea of the definitional approach. The mathematician sees that he needs a new number in J, call it 1^*, such that when combined with the number 1 in J under the new operation $+_J$ it gives 0; that is,

$$1 +_J 1^* = 0$$

We now look through the elements of J seeking a number y such that $2 +_J y = 0$

Is there one?

Lecture 14, May 1, 1968.

We have been studying the whole number system $(W,+,\cdot)$, which for various reasons already discussed, we want to extend to a new number system $(J,+_J,\cdot_J)$; this extension means: finding a set J which has W as a subset and finding operations $+_J$ and $\cdot_J$ on this set J such that whenever x,y are elements of W, then $x +_J y = x + y$ and $x \cdot_J y = x \cdot y$ In order that this number system $(J,+_J,\cdot_J)$ be indeed different from $(W,+,\cdot)$ we also require that J have a number x in it such that $1 +_J x = 0$.

Question I: Can we find such a system $(J,+_J,\cdot_J)$ containing only one new number in J in addition to the old numbers of W? We can answer this in the following way: suppose 1^* is a number in J such that $1 +_J 1^* = 0$. Since J is closed under $+_J$ we must have that $1^* +_J 1^*$ is also in J. Since we are assuming here that $J = W \cup \{1^*\}$, it must be that $1^* +_J 1^*$ is a number in this set. Which one can it be?

Could it be that $1^* +_J 1^* = 0$?
No, because if $1^* +_J 1^* = 0$, then the following would hold:

$$1^* +_J 1^* = 0 = 1 +_J 1^*$$

$$\text{i.e. } 1^* +_J 1^* = 1 +_J 1^*$$

and using the cancellation law, we would get

$$1^* = 1$$

This is false (i.e. $1^* \neq 1$) since 1 is in W and 1^* is not in W. So by assuming that $1^* +_J 1^* = 0$, we are led to a contradiction. We thus know that $1^* +_J 1^* \neq 0$

Could it be that $1^* +_J 1^* = 1$?
No, because if $1^* +_J 1^* = 1$, then by adding 1 to both sides we would get

$$1 +_J (1^* +_J 1^*) = 1 + 1$$

and using the associative law, we would get

$$(1 +_J 1^*) +_J 1^* = 1 + 1$$

but we know $1 +_J 1^* = 0$, so the above equation reduces to:

$$0 +_J 1^* = 2,$$

which by the identity law for $+_J$ reduces to $1^* = 2$
This also is false since 1^* is not in W, whereas 2 is. We have again been led to a contradiction, so our hypothesis again must also be false--that is, we now know that

$$1^* +_J 1^* \neq 1$$

You should now see the pattern and be able to show that

$$\begin{aligned} 1^* +_J 1^* &\neq 2 \\ &\neq 3 \\ &\neq 4 \\ &\vdots \quad \vdots \end{aligned}$$

So we know that $1^* +_J 1^*$ is not equal to any whole number.
But

Could it be that $1^* +_J 1^* = 1^*$?
NO, because if this were so; that is, if $1^* +_J 1^* = 1^*$, then adding $\overline{1}$ to both sides of this equation, we get

$$1 +_J (1^* +_J 1^*) = 1 + 1^*$$

$$\text{i.e. } (1 +_J 1^*) +_J 1^* = 0$$

$$\text{i.e. } 0 +_J 1^* = 0$$

$$\text{i.e. } 1^* = 0,$$

but in fact this is not so, since 0 is a whole number and 1^* is not a whole number. Thus, we've answered the above Question in the negative, by showing $1^* +_J 1^*$ is not a whole number, and in addition it is not equal to 1^*. Hence $1^* + 1^*$ must be a second new number of J. What else can we say about this second new number $1^* +_J 1^*$ in J?

We claim: $2 +_J (1^* +_J 1^*) = 0$

Proof: $2 = 1 + 1$

$$\text{hence, } \quad 2 +_J (1^* +_J 1^*) = (1 + 1) +_J (1^* +_J 1^*)$$
$$= (1 +_J 1) +_J (1^* +_J 1^*)$$

and by two applications of the associative law for $+_J$,

$$= 1 +_J ((1 +_J 1^*) +_J 1^*),$$

which by definition of 1^* $= 1 +_J (0 +_J 1^*)$,

which by the commutative and identity laws for $+_J$ $\Big\}$ $= 1 +_J 1^*$

$$= 0$$

Thus, $2 +_J (1^* +_J 1^*) = 0$ and knowing this it is natural to introduce the

Definition: $2^* = 1^* + 1^*$

Now, using this definition and the above claim which we have just proved, we get that $2 +_J 2^* = 0$

Thus, we have found our new system must contain at least two new numbers, 1^* and 2^*.

Question II: Do we now have enough new numbers to satisfy our requirements? No, we don't, and we would show this by considering the number $1^* +_J 2^*$, which must be in J because J is required to be closed under $+_J$. Arguing as before we could show:

$$\begin{aligned} 1^* +_J 2^* &\neq 0 \\ &\neq 1 \\ &\neq 2 \\ &\;\vdots\;\;\vdots \\ &\neq 1^* \\ &\neq 2^* \end{aligned}$$

That is, $1^* +_J 2^*$ is a third new number. Moreover, we could show that $3 +_J (1^* +_J 2^*) = 0$, making it natural to introduce the

Definition: $3^* = 1^* +_J 2^*$,

so that we would have $3 +_J 3^* = 0$

Proceeding in this way and using the Principle of Mathematical Induction to obtain full generality, we would finally see that: for every old number z, except 0, there must be a new number z^* in J such that $z +_J z^* = 0$

Furthermore, if x,y are two different old numbers each different from 0, then x^* and y^* will also be different from each other.

We have been considering the following Problem. We wish to find a number system $(J, +_J, \cdot_J)$ which

(i) is an extension of $(W, +, \cdot)$
(ii) satisfies laws similar to those holding in the system of whole numbers, e.g. the commutative, associative, cancellation, and identity laws for $+_J$, similar laws for $\cdot_J$, etc., and
(iii) contains a number x such that $1 +_J x = 0$

We have already found that if we have a system satisfying (i), (ii), (iii), then for every z in W, $z \neq 0$, there must be an element z^* in J such that

(a) $z + z^* = 0$, and
(b) all these elements z^* are new numbers, i.e., they are not whole numbers of W.

Furthermore, if x,y are whole numbers $\neq 0$ and if $x \neq y$, then (c) $x^* \neq y^*$.

Thus J must contain infinitely many distinct new numbers (i.e., numbers not in W): $1^*, 2^*, 3^*, \ldots$.

Now that we know something about the size of J, let us find out something about how the operation $\cdot_J$ must work.

First of all we have

(A) $1^* \cdot_J 0 = 0$ and $1^* \cdot_J 1 = 1^*$

because of the general laws $x \cdot_J 0 = 0$ and $x \cdot_J 1 = x$ which are among the desiderata (ii) above. Using the second equation of line (A) above, together with the distributive law, we find

$$\begin{aligned} 1^* \cdot_J 2 = 1^* \cdot_J (1 +_J 1) &= (1^* \cdot_J 1) +_J (1^* \cdot_J 1) \\ &= 1^* +_J 1^* \\ &= 2^* \end{aligned}$$

by definition of 2^*.

Similarly, we can find $1^* \cdot_J 3 = 3^*$, $1^* \cdot_J 4 = 4^*$, ... , and more generally, for every z in W (other than 0),

(B) $1^* \cdot_J z = z^*$.

Next let us compute $1^* \cdot_J 1^*$. Using the fact that $1 +_J 1^* = 0$, from (a) above, and the distributive law, we find

$1^* \cdot_J (1 +_J 1^*) = 1^* \cdot_J 0$

$= 0$ from line (A)

$(1^* \cdot_J 1) +_J (1^* \cdot_J 1^*) = 0$ (distributive law)

$1^* +_J (1^* \cdot_J 1^*) = 1 +_J 1^*$ (by (A) and (a))

$= 1^* +_J 1$ by commutative law (ii)

(C) Hence $1^* \cdot_J 1^* = 1$, by cancellation law for $+_J$.

Now if z is any whole number $\neq 0$, then

$1^* \cdot_J z^* = 1^* \cdot_J (1^* \cdot_J z)$ by line (B) above

$= (1^* \cdot_J 1^*) \cdot_J z$ by associative law for $\cdot_J$

$= 1 \cdot_J z$ by line (C) above

$= z$ by identity law for $\cdot_J$.

(D) Thus we've shown: $1^* \cdot_J z^* = z$, for z in W, $z \neq 0$.

Next, if y, z in W and $y, z \neq 0$ then

$y^* \cdot_J z^* = (y \cdot_J 1^*) \cdot_J z^*$ by line (B) and commutative law for $\cdot_J$

$= y \cdot_J (1^* \cdot_J z^*)$ by associative law, $\cdot_J$

$= y \cdot_J z$ by line (D) above

$= y \cdot z$ since $\cdot_J$ reduces to $\cdot$ when acting on numbers in W.

(E) Thus we've shown $y^* \cdot_J z^* = y \cdot z$ whenever y, z in W and $y, z \neq 0$.

Finally, if y, z in W and $y, z \neq 0$ then

$y \cdot_J z^* = z^* \cdot_J y$ by commutative law, $\cdot_J$

$= (1^* \cdot_J z) \cdot_J y$ by (B)

$= 1^* \cdot_J (z \cdot_J y)$ by associative law, $\cdot_J$

$= 1^* \cdot_J (z \cdot y)$ since $\cdot_J$ reduces to $\cdot$ on W

$= (z \cdot y)^*$ by (B) again.

(F) We've shown $y \cdot_J z^* = (z \cdot y)^*$ and

$z^* \cdot_J y = (z \cdot y)^*$

whenever y, z in W and $y, z \neq 0$.

By combining (A), (E), (F) we get a complete rule for carrying out the operation $\cdot_J$ on any pair of numbers of J, new or old. Similarly, one can find how to carry out the operation $+_J$.

With all this knowledge about how $(J, +_J, \cdot_J)$ must look if it is to satisfy (i), (ii), (iii) above, we can (in various ways) construct such a system.

For instance, given any z in W, $z \neq 0$, we can define z^* to be the ordered pair $(z, \text{Antares})$. These "numbers" z^* are not in W, and if $y \neq z$ then $y^* \neq z^*$, so we can define J to be the set $W \cup \{1^*, 2^*, 3^*, \ldots\}$ and then we can define $\cdot_J$ by the formulas (A), (E), (F) above, and we can define $+_J$ by analogous formulas.

It is then possible to prove that the system $(J, +_J, \cdot_J)$ so defined satisfies (i)-(iii).

Lecture 15, May 6, 1968.

As a reminder of our notation, recall that we denoted the set of whole numbers by: $W = \{0,1,2,3,\ldots\}$

zero by: 0

and the Natural numbers (or counting numbers) by: $N = \{1,2,3,\ldots\}$

Also, we have begun to study the set of Negative numbers: $\{1^*,2^*,3^*,\ldots\}$, where, for each natural number b, we defined $b^* = (b, \text{Antares})$. Putting these negative numbers into a set together with the whole numbers, we get the set J which we called the set of Integers. i.e. we defined the set J as follows: $J = W \cup$ The set of all negative numbers.

$$= \left\{\begin{array}{l} 0,1,2,3,\ldots, \\ 1^*,2^*,3^*,\ldots \end{array}\right\}$$

On this new set J we defined the operation multiplication, written $\cdot_J$, as follows:

(i) For two natural numbers

If b,c are any natural numbers, $b \cdot_J c = b \cdot c$

(ii) For two negative numbers

If b,c are any natural numbers, $b^* \cdot_J c^* = b \cdot c$

(iii) For one natural number and one negative number

If b,c are any natural numbers, $b \cdot_J c^* = (b \cdot c)^*$ and

$c^* \cdot_J b = (c \cdot b)^*$

(iv) If one number is zero

Given any integer x, $0 \cdot_J x = 0$ and

$x \cdot_J 0 = 0$

Also, we have the Definition of $+_J$ on the set J

(i) For two natural numbers

If a,b are any natural numbers, $a +_J b = a + b$

(ii) For two negative numbers

If a,b are any natural numbers, $a^* +_J b^* = (a + b)^*$

(iii) For one natural number and one negative number

If a,b are any natural numbers,

$$a +_J b^* = \begin{cases} a - b, & \text{if } a \geq b \\ (b - a)^* & \text{if } a < b \end{cases}$$

$$b^* +_J a = \begin{cases} (b - a)^* & \text{if } b > a \\ a - b & \text{if } b \leq a \end{cases}$$

(iv) If one number is zero

If x is any integer, $0 +_J x = x$

$$x +_J 0 = x$$

By way of examples,
Notice, from part (iii) of the definition of $+_J$, that

a) $1^* +_J 1 = 1 - 1 = 0$, and
b) $1 + 1^* = 1 - 1 = 0$
c) $3 +_J 8^* = (8 - 3)^* = 5^*$
d) $8 +_J 3^* = 8 - 3 = 5$

And from part (iii) of the definition of $\cdot_J$, that

$$3 \cdot_J 8^* = (3 \cdot 8)^* = 24^*$$

We have now completed our definition of the Integer Number System $(J, +_J, \cdot_J)$: It can be proven that this new system satisfies the three conditions we previously desired any extension of $(W, +, \cdot)$ to satisfy, although we will not do so here. It should be immediately clear, however, that $(J, +_J, \cdot_J)$ is an extension of $(W, +, \cdot)$.

Definition: We introduce on J a one-place (unary) operation, called negation and denoted by $^{-}$; that is, given any integer z of J, we apply the operation to get another integer ^{-}z, which we define as follows: If a is any natural number, then

$$^{-}a = a^*$$

$$^{-}(a^*) = a$$

$$^{-}0 = 0$$

Question: For x an integer, is ^{-}x a negative number? You can't tell yet, since we don't have enough information about x. However, we can say that:

If x is a positive number, ^{-}x is a negative number.
If x is a negative number, ^{-}x is a positive number.
If x is zero, ^{-}x is zero.

There are some general laws involving this new operation negation, which you should be aware of:

a) For all integers z, $^{-}(^{-}z) = z$
Note that this follows from the definition of the operation negation.

For all integers x,y

b) $(^{-}x) \cdot_J (^{-}y) = x \cdot_J y$

c) $(^{-}x) +_J (^{-}y) = {}^{-}(x +_J y)$

d) $(^{-}x) \cdot_J (y) = {}^{-}(x \cdot_J y)$

Let's go through a demonstration of (b) above; i.e. we'll prove $(^{-}x) \cdot_J (^{-}y) = x \cdot_J y$, by using the definition of negation and $\cdot_J$.

case i: If x,y are natural numbers,
then $(^{-}x) \cdot_J (^{-}y) = (x^*) \cdot_J (y^*) = x \cdot y = x \cdot_J y$

case ii: If x,y are both negative numbers,
then $x = a^*$ and $y = b^*$, where a,b are natural numbers so

$$x \cdot_J y = (a^*) \cdot_J (b^*) = a \cdot b$$

also,

$$\begin{aligned}(^{-}x) \cdot_J (^{-}y) &= {}^{-}(a^*) \cdot_J {}^{-}(b^*) \\ &= a \cdot_J b \\ &= a \cdot b\end{aligned}$$

Thus, $(^{-}x) \cdot_J (^{-}y) = x \cdot_J y$

case iii: If x is a natural number and y a negative number, then $y = b^*$, where b is a natural number. So $(^-x) \bullet_J (^-y) = (^-x) \circ_J {}^-(b^*)$

$$= x^* \bullet_J b$$

$$= (x \circ_J b)^*$$

Also, $(x \bullet_J y) = x \circ_J b^* = (x \circ_J b)^*$

Thus, $(^-x) \circ_J (^-y) = x \bullet_J y$

case iv: If either x or y is zero.
With no loss of generality, suppose $x = 0$ and y is any integer.
Then $(^-x) \bullet_J (^-y) = (^-0) \bullet_J (^-y)$

$$= 0 \bullet_J (^-y)$$

$$= 0$$

Also, $x \bullet_J y = 0 \circ_J y = 0$

Hence, $(^-x) \bullet_J (^-y) = x \bullet_J y$

The other three laws (a), (c), (d) above may be proven similarly by going through the four cases as we did in proving (b).

Lecture 16, May 8, 1968.

Today we will finish our study of the set of integers

$J = \{\underbrace{1,2,3,\ldots}_{\text{natural numbers}}, \underbrace{0}_{\text{zero}}, \underbrace{1^*,2^*,3^*,\ldots}_{\text{negative integers}}\}$.

Although we have found that for all integers x,y $(^-x) \cdot_J (^-y) = x \cdot_J y$, this does not say that the product of two negative integers is a positive integer, since (^-x) may in fact be positive. In order to express the general proposition that a product of two negative integers is a positive integer, we would write: For all natural numbers, b,c, $b^* \cdot_J c^* = b \cdot c$.

Finally on J, we want to introduce the operation ^-J, called subtraction. Recall that on the set W, subtraction had many messy conditions accompanying its use. This will not be the case on J; here, precisely is one of the reasons we decided to extend the set W.

First we need the

Theorem: Given any integers x,y, there is one and only one integer z such that $x +_J z = y$.

Once, this theorem is established we can introduce the

Definition of Subtraction on J:

Given any integers x,y, we define $y -_J x$ to be the unique integer z such that $x +_J z = y$.

How we would prove the above theorem: There are many cases to consider but once you see the general procedure, it will be easy for you to complete the proof.

Case 1: Suppose x, y are both natural numbers.

Subcase 1.a: Suppose $y \geq x$. Then by the theory of subtraction for the whole numbers, there is a whole number z (namely, $y - x$) such that $y = x + z$. This whole number z is an integer because all the whole numbers are among the integers and also $y = x +_J z$, since $+_J$ is an extension of $+$.

Subcase 1.b: Suppose not $y \geq x$; i.e. it is the case that $y < x$. Then since there is no whole number which can be added to x to give us y, we must look for a negative number which will work. Let's take a whole number b (namely, $x - y$) such that $y + b = x$. Furthermore, $b \neq 0$, since $y + 0 = y$ and $y + b = x$, and $x \neq y$, since by hypothesis $y < x$.

Thus, we know b is a whole number different from zero—that is, a natural number. So there is a negative number b^*. Let's see if b^* works--that is, we would like to show: $x +_J b^* = y$.

We know: $y +_J b = x$

Adding b^* to both sides using the logic of equality, we get $(y +_J b) +_J b^* = x +_J b^*$.

The associative law for $+_J$, applied to the left side, gives us
$$y +_J (b +_J b^*) = x +_J b^*$$
i.e.
$$y +_J 0 = x +_J b^*$$
and by the additive identity for $+_J$, $y = x +_J b^*$

This completes the proof of Case 1; the other cases follow analogously.

We finish with a few examples of general statements for the subtraction operation $-_J$.

(i) For all integers x, $x -_J 0 = x$ and

$$x -_J x = 0.$$

(ii) For all integers x, y, z,

$$x -_J (y +_J z) = (x -_J y) -_J z$$

Note: no messy conditions here, as there were when the analogous law for W was discussed.

(iii) For all integers x, $0 -_J x = {}^{-}x$

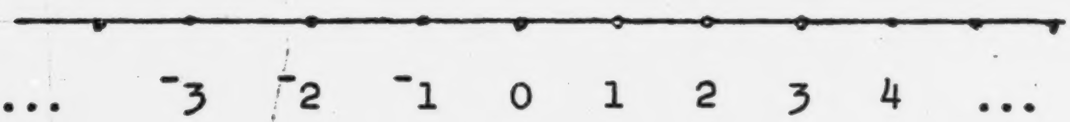
Notice that x may be a negative number, zero, or a natural number.

And Most Importantly,

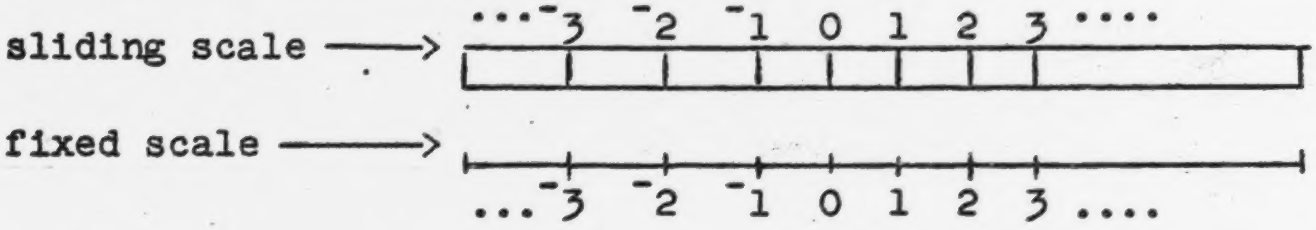
(iv) For all integers x, y, $x +_J ({}^{-}y) = x -_J y$

Now, with the idea of making addition, subtraction and multiplication more intuitive, we'll discuss an application of the negative numbers, one connecting the negative numbers with geometry. We'll consider a straight line, straight in the sense of Euclidean geometry; in principle, it has no left end and no right end. We distinguish a line (infinite) from a line segment which has ends. Arbitrarily, we place a point on this line we call 0, and to the right of 0 we arbitrarily place another point we call 1. Example:

We'll now use the distance from 0 to 1 as our unit for measuring length, and using it we can proceed to label the points 2,3,4,... to the right of 1 and the points $^{-}1, ^{-}2, ^{-}3, \ldots$ to the left of 0. This line with the points so labeled is called in elementary school a number line; a partial representation of it would look like this:

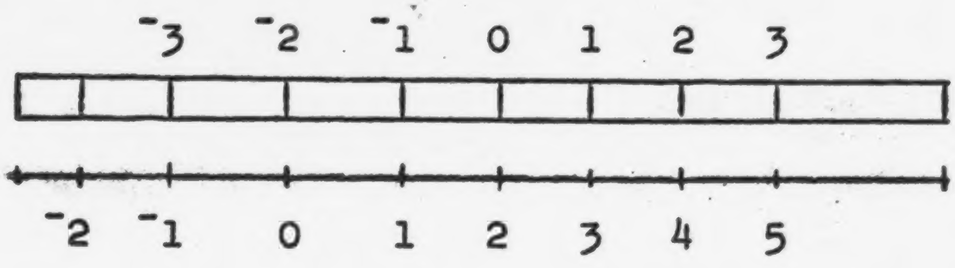


We'll now construct a kind of slide rule. (In fact, you could call it an analogue computer.) Imagine it to be made of two pieces of some material, one piece sliding on the other; both infinite, but one is fixed while the other can slide next to it. On each is a representation of a number line. It looks about like this:



Here is the way it works: Suppose we want to add the two integers 2 and $^{-}3$. That is, we want to find $2 +_J {}^{-}3$.

(1) Take the 0 point on the sliding scale and slide it over to the number 2 on the fixed scale, like this:



(ii) Find ${}^{-}3$ on the sliding scale and

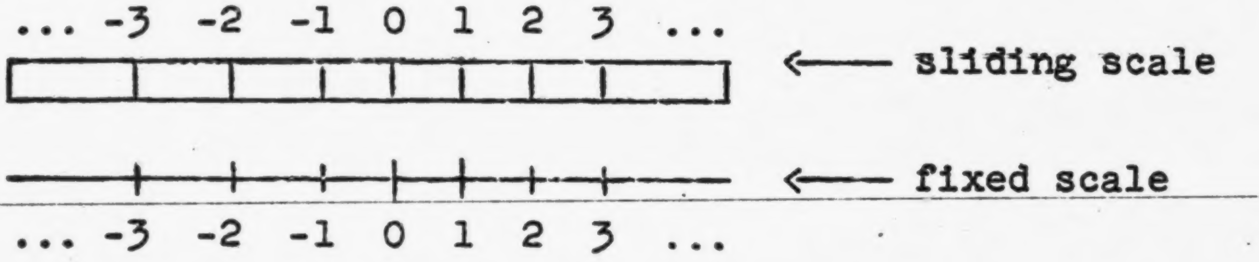
(iii) Obtain the answer by reading the number opposite this on the fixed scale, in this case ${}^{-}1$.

Thus $2 +_J {}^{-}3 = {}^{-}1$.

Finally, let us mention that our use of $+_J$ and $\cdot_J$ are pedagogical devices, helpful for understanding the extension of W to J. Once we have the knowledge of these new numbers and the laws governing them, we never need to return to our old definitions. Hence, from now on, we'll drop the subscript J on + and $\cdot$.

Lecture 17, May 10, 1968.

As an intuitive aid to understanding addition in the set J of integers, we introduced a hypothetical slide rule:



To find the sum, $x + y$, where x and y are any two given integers,

(i) Move the 0 point on the sliding scale opposite x on the fixed scale.

(ii) Find y on the sliding scale, and

(iii) The number on the fixed scale opposite y on the sliding scale is the desired number $x + y$.

Notice, from the side rule, we can distinguish the following:

Case 1: If x, y are natural numbers, then so is $x + y$.
Example: $1 + 4 = 5$

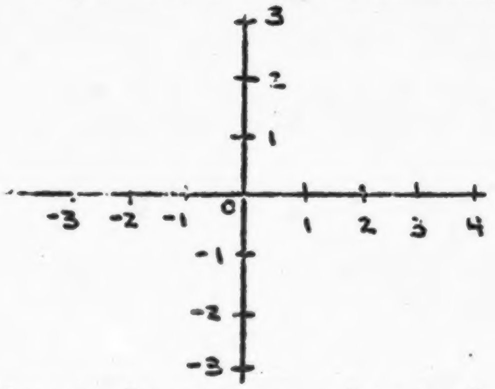
Case 2: If x, y are negative numbers, then so is $x + y$, since the scale is moved to the left.
Example: ${}^{-}3 + {}^{-}2 = {}^{-}5$

Case 3: If x is positive and y is negative (or vice-versa) then the nature of $x + y$ depends on the values of x and y.
Examples: ${}^{-}3 + 2 = {}^{-}1$
$1 + {}^{-}4 = {}^{-}3$
${}^{-}2 + 3 = 1$

Furthermore, we can use our slide rule for the operation subtraction, if we first recall the most important general law about subtraction on J which stated: For any integers x, y, $x - y = x + {}^{-}y$. That is, we rewrite the subtraction problem as an addition. Example: ${}^{-}4 - {}^{-}6$, which by above is the same as ${}^{-}4 + {}^{-}({}^{-}6)$, which is the same as ${}^{-}4 + 6$, because for every integer x, ${}^{-}({}^{-}x) = x$.

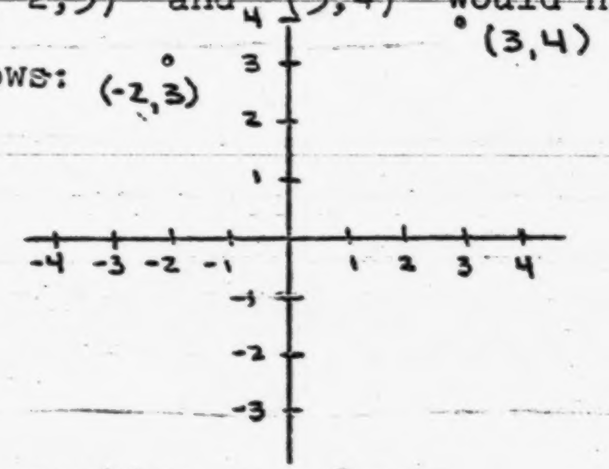
Thus, we are left with a simple sum of two integers which can be carried out on the slide rule as before.

Although the slide rule is an aid for addition and subtraction of integers, for multiplication we introduce the concept of graphs. Suppose we have an equation such as $y = 2 \cdot x$; associated with this equation is a picture called a graph derived as follows: First we draw two number lines, one horizontal, called the horizontal or x-axis and one vertical, the vertical or y-axis; both of these lines crossing at their respective 0 points, as follows:



A pair of number lines drawn as above is called a set of co-ordinate axes. Within the plane of these co-ordinate axes, we place a point, the geometrical picture of any ordered pair of integers (x,y), where the first member of the ordered pair refers to the distance and direction along the horizontal axis and the second member refers to the distance and direction along the vertical axis.

The two ordered pairs $(^-2,3)$ and $(3,4)$ would have their representation as follows:



Now, to get the graph of the equation $y = 2 \cdot x$,

(i) Choose various arbitrary integers as values for x.

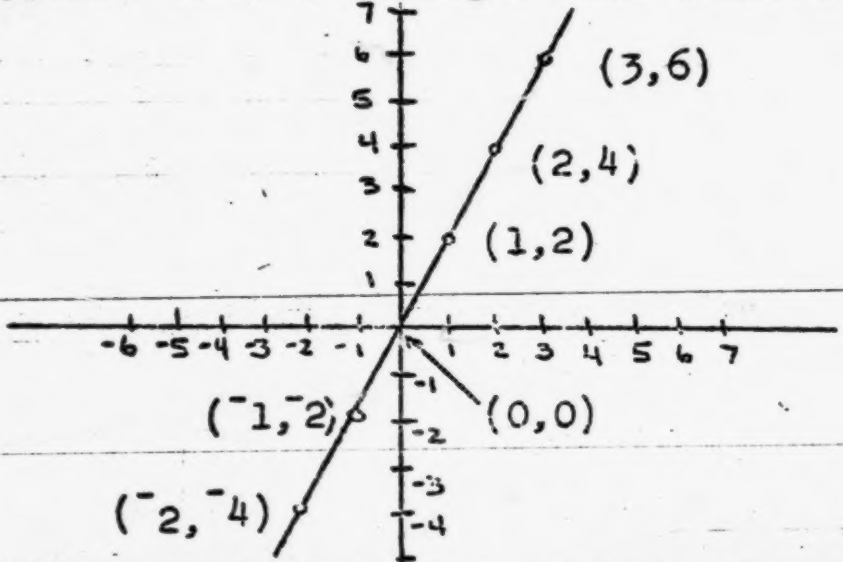
(ii) Compute the corresponding value for y, and then

(iii) Plot (draw) the picture of the ordered pairs (x,y).

Using the equation $y = 2 \cdot x$, we first choose values of x, say $0,1,2,3,^-1,^-5$; these are arbitrary. Now we compute, using the given equation, the values of y corresponding to these values of x and conveniently list them in a table:

x	y
0	0
1	2
2	4
3	6
$^-1$	$^-2$
$^-2$	$^-4$

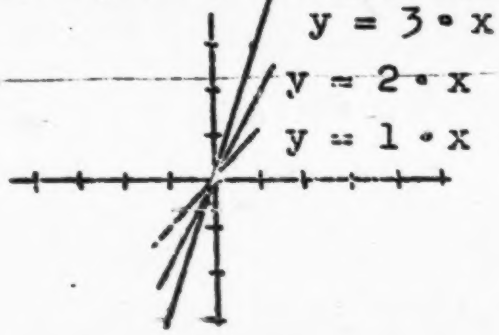
Using these five ordered pairs $[(0,0),(1,2),(2,4),(3,6),(^-1,^-2)$ $(^-2,^-4)]$, we plot the graph of these points drawing in what appears to be a straight line connecting them, as follows:



Graph of
$y = 2 \cdot x$

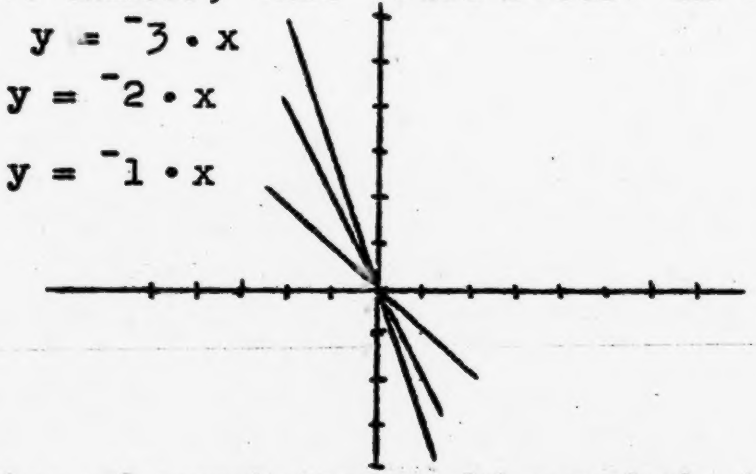
Notice, that if we did not know the product of 2 and ¯1 or the product of 2 and ¯2, the graph of $y = 2 \cdot x$ could suggest them, since even without these products, we would have a straight line connecting (3,6), (2,4) and (1,2). Extending this line on both sides, to arrive at $2 \cdot {}^{-}1$, we would locate ¯1 on the horizontal axis and notice that the ordered pair on the line of the graph of $y = 2 \cdot x$ whose first co-ordinate is ¯1 is (¯1,¯2). Thus $2 \cdot {}^{-}1 = {}^{-}2$. Similarly to arrive at $2 \cdot {}^{-}2$, we first locate ¯2 on the horizontal axis and see that the ordered pair on the graph of $y = 2 \cdot x$ whose first co-ordinate is ¯2 is (¯2,¯4). Thus $2 \cdot {}^{-}2 = {}^{-}4$.

Now drawing the graphs of $y = 1x$, $y = 2 \cdot x$, $y = 3 \cdot x$, etc., you can see that as the whole number which we multiply x by increases, the corresponding graphs increase in steepness.



From these graphs, this much should be clear: If the "x-value" is negative, the corresponding ordered pair on any of these graphs will have a negative "y-value". Thus, we can conclude that when we multiply a negative number by the positive number 1, 2 or 3, the result is negative. Moreover, we could continue drawing the graphs of $y = b \cdot x$ for b any positive integer, getting steeper and steeper graphs, and then be able to conclude that the product of a positive and a negative number will always be negative.

Once this is done, we could plot the graphs of $y = {}^-1 \cdot x$, $y = {}^-2 \cdot x$, $y = {}^-3 \cdot x$, or in general $y = {}^-b \cdot x$, where b is any positive number; this would result in the following graphs:



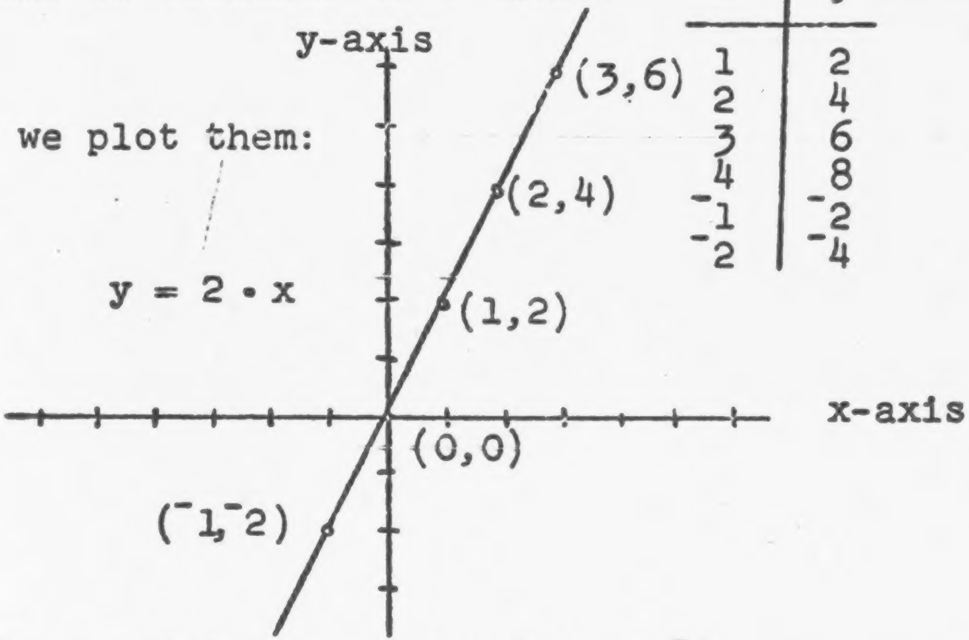
From this series of graphs we could conclude that for x any negative number, the corresponding value for y in the graph would be positive--that is, the product of any two negative numbers in a positive number.

Lecture 18, May 13, 1968.

Let us quickly review some of the details of graphing using as our first example the equation $y = 2 \cdot x$. We first arbitrarily select values for x substituting them into this equation and so getting a value for y. We summarize the values so obtained in a table:

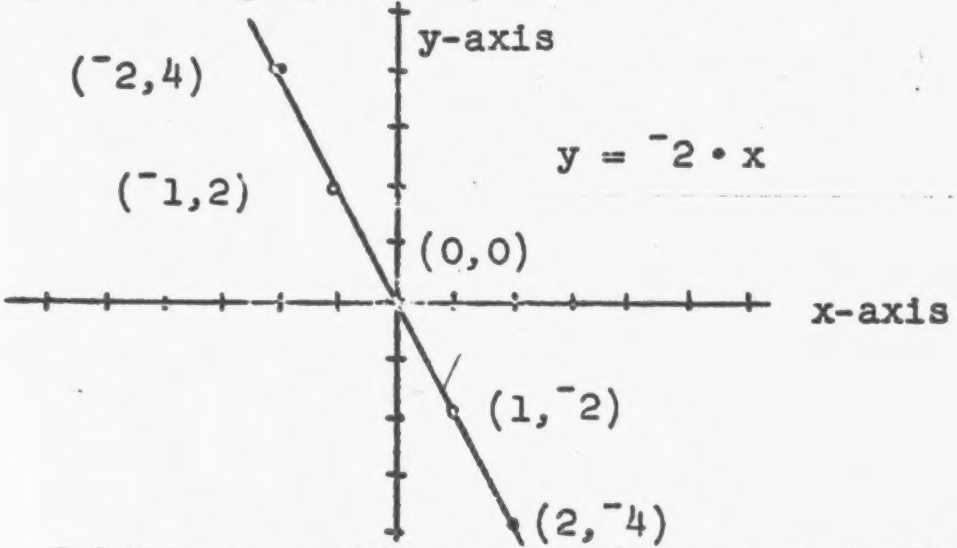
x	y
1	2
2	4
3	6
4	8
$^{-}1$	$^{-}2$
$^{-}2$	$^{-}4$

And now we plot them:



Similarly we plot the graph of $y = {}^{-}2 \cdot x$

x	y
1	$^{-}2$
2	$^{-}4$
3	$^{-}6$
4	$^{-}8$

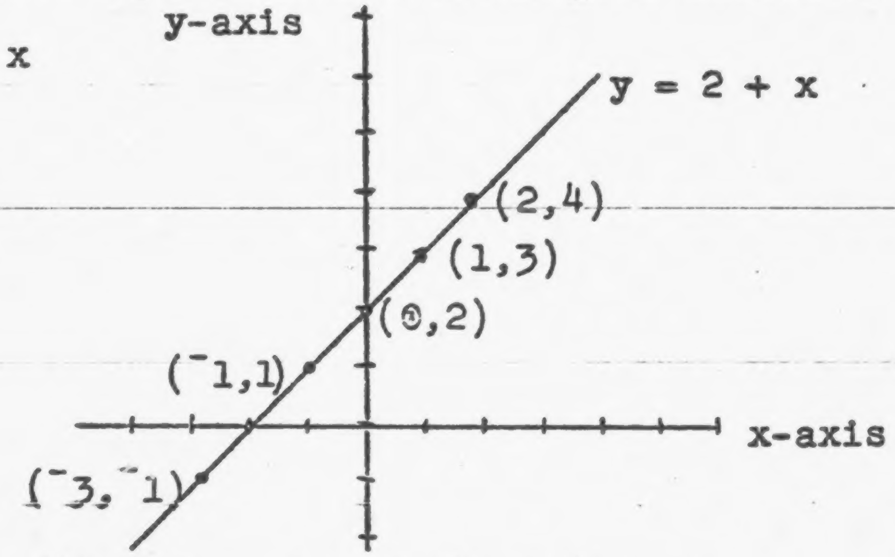


Notice that: Taking a negative value for x and asking what is the corresponding value of y (looking only at the graph) we see that the product of two negative numbers gives us a positive number .

Finally, we mention that graphs of additive equations as well as the multiplicative equations are meaningful entities..

Example: $y = 2 + x$

x	y
0	2
1	3
2	4
3	5



Some Other Aspects of the Integers

If we're given a positive integer, say 5, then its clearly possible to break it into a sum of two smaller integers; for example, $5 = 2 + 3$. Similarly we can break 2 and 3 down, finally getting $5 = 1 + 1 + 1 + 1 + 1$. It should also be quite clear that every positive integer may be broken down to a sum of 1's.

However, for the operation multiplication, the case is quite different. For example, 5 cannot be broken down into a product of two (or more) smaller integers, and so 5 is called a prime number. As another example, 196 can be broken down: first, since 196 is an even number, we see that $196 = 2 \cdot 98$. However, we can go still further.

i.e.
$$\begin{aligned} 196 &= 2 \cdot 98 \\ &= 2 \cdot 2 \cdot 49 \\ &= 2 \cdot 2 \cdot 7 \cdot 7 \\ &= 2^2 \cdot 7^2 , \end{aligned}$$
using our exponential notation.

Notice, we cannot continue this process further since 2 and 7 cannot be broken down into a product of smaller positive integers; they are both prime numbers. Thus, we have represented 196 as the product of two prime numbers each "taken" twice. In fact, we have the

Theorem Every integer not 0, $^{-}1$, or 1, which is not itself prime can be expressed as a product of positive prime numbers (with a factor $^{-}1$ in case the given number is negative) and indeed in only one way (aside from different orderings of the prime factors or extra factors of 1).

Consequently, in the set J of all integers, we distinguish the following types of numbers:

(i) *0* : The additive identity element.

(ii) *Units* : 1, $^{-}1$. These elements have a multiplicative inverse in the set J--i.e. given either of these number, say $^{-}1$, we can find a number in J which when multiplied by the given number results in the multiplicative identity element in J, which is 1. In other words, $^{-}1$ is its own multiplicative inverse, and 1 is its own multiplicative inverse since $^{-}1 \cdot {}^{-}1 = 1$ and $1 \cdot 1 = 1$.

Notice, however that 2 has *no* multiplicative inverse; that is, there is no integer which when multiplied by 2 gives us 1. In fact, there are no integers except $^{-}1$ and 1 which have multiplicative inverses in the set J. Thus the only units in the set J are 1 and $^{-}1$.

Before we continue, we introduce the

Definition of Divisor: Given an integer x, we call z a divisor of x or a factor of x just in case there is some integer w such that $x = z \cdot w$.

Now, all other elements in J are divided into two kinds:

(iii) Primes : These are the integers x, other than 0 and the units, whose only factors are 1, ¯1, x, ¯x. For example, 5 is a prime, since the only divisors of 5 are 1, ¯1, 5, ¯5. Similarly, ¯13 is a prime.

(iv) Composites : All integers other than 0, the units and the primes.
Example: ¯26 is a composite since it is not 0, not a unit and not a prime. (¯26 = 2 • ¯13)

Now that we have subdivided the integers into various categories, lets investigate a few applications. First, a

Definition : Given any positive integers x,y, by a common divisor of x and y, we mean an integer z which is a divisor of both x and y.

We now state (but do not prove) a

Theorem: Among all common divisors of x and y there is a certain positive integer z such that every other common divisor of x and y is also a divisor of z. Thus we have the

Definition: This unique common divisor of x and y is called the greatest common divisor of x and y, abbreviated g.c.d.(x,y).

Examples: (1) Given the two integers 12 and 18, we first express them as a product of prime numbers according to our theorem. $12 = 2 \cdot 2 \cdot 3$

Now to find all the positive factors of 12, immediately we know 1 is, and then we find the products of its prime factors; first, one at a time; then two at a time, then three at a time. Thus the positive divisors of 12 are: 1, 2, 3, 4, 6, 12. Similarly for $18 = 2 \cdot 9 = 2 \cdot 3 \cdot 3$. Its divisors are: 1,2,3,6,9,18. Now the common divisors of 12 and 18 are: 1, 2, 3, 6 and the g.c.d.(12,18) is 6, since all the other common divisors of 12 and 18 are divisors of 6.

Lecture 19, May 17, 1968.

At the last lecture, we introduced the following two definitions:

1) *Definition of Divisor*. An integer x is a *divisor* or *factor* of an integer y just in case there is an integer z such that $x \cdot z = y$. To express this we use the notation $x|y$, which is read x divides y or if it is *not* the case that x divides y, then we write $x \nmid y$.

   For example, $3|6$, $3|18$, $3 \nmid 20$.

2) *Definition of Prime*. An integer y is a *prime number* just in case the only factors of y are $1, -1, y, -y$.

   Examples: 7 is a prime

   14 is not a prime

Concerning these newly defined notions of prime and divisor, we now state a few *General Statements* or *Laws*:

For divisor we have: 1) *Transitive Law for Divides*

If x, y, z are any integers such that $x|y$ and $y|z$, then also $x|z$. In order to prove this law, which we leave as an exercise, look carefully at the proof previously given for the transitive law for the relation $<$.

2) If x and y are any integers such that $x|y$ and $y|x$ then $x = y$ or $x = {}^{-}y$.

For Primes we have: 1) If x is *any* positive integer, then there is a prime number z such that $z > x$. This famous theorem involving primes goes back to Euclid and his proof of it is both simple and elegant.[1] Notice that the law states that there are infinitely many prime numbers.

2) If z is any positive even number, then there are numbers x and y which are either prime or 1 such that $z = x + y$. This is a general statement since it involves all positive even numbers. It was first conjectured by a man named Goldbach about 200 years ago and it still remains a conjecture today--that is, its truth or falsity has not been determined.

We have previously introduced the definition of *greatest common divisor* of a given pair of positive integers x and y. There are two basic methods for finding this g.c.d.(x,y).

Method 1: Factor x and then y into their prime factors. Say $x = p_1^{a_1} \cdot p_2^{a_2} \cdot \ldots \cdot p_n^{a_n}$ and $y = p_1^{b_1} \cdot p_2^{b_2} \cdot \ldots \cdot p_n^{b_n}$, where $p_1, p_2, \ldots, p_n$ are the prime factors of x and y, and $a_1, a_2, \ldots, a_n, b_1, b_2, \ldots, b_n$ are the whole number exponents of these factors. Some of the a's and b's may be zero. Then the g.c.d.(x,y) is the number z, where $z = p_1^{c_1} \cdot p_2^{c_2} \cdot \ldots \cdot p_n^{c_n}$ and where

---

[1] See *What is Mathematics?* by Courant and Robbins.

c_1 = minimum of (a_1, b_1)

c_2 = " " (a_2, b_2)

$\vdots$

c_n = " " (a_n, b_n)

Notice: minimum of $(a_1, b_1) = \begin{cases} a_1 & \text{if } a_1 \leq b_1 \\ \text{or} & \\ b_1 & \text{if } b_1 \leq a_1 \end{cases}$

For example: To find g.c.d.(12,18), first we factor 12 and 18 into their prime factors.

$$12 = 2 \cdot 6 = 2 \cdot 2 \cdot 3 \qquad \text{i.e.} \quad 12 = 2^2 \cdot 3^1$$

$$18 = 2 \cdot 9 = 2 \cdot 3 \cdot 3 \qquad \text{i.e.} \quad 18 = 2^1 \cdot 3^2$$

Therefore $\text{g.c.d.}(12,18) = 2^{\min(2,1)} \cdot 3^{\min(1,2)}$

$$= 2^1 \cdot 3^1 = 6$$

So g.c.d.(12,18) = 6.

It is a lucky accident that the primes in the above two decompositions are the same. The following example shows that regardless of the original prime decomposition, we can introduce primes raised to the zero power in the decomposition so that they will always be the same.

Say, we want to find g.c.d.(48,76). We first factor the two numbers.

$$48 = 2 \cdot 2 \cdot 4 \cdot 3 = 2 \cdot 2 \cdot 2 \cdot 2 \cdot 3 = 2^4 \cdot 3$$

$$76 = 2 \cdot 38 = 2 \cdot 2 \cdot 19 = 2^2 \cdot 19$$

But also, $48 = 2^4 \cdot 3^1 \cdot 19^0$

$$76 = 2^2 \cdot 3^0 \cdot 19^1$$

Lecture 19, May 17, 1968.

Hence, $\text{g.c.d.}\ (48,76) = 2^{\min(4,2)} \cdot 3^{\min(1,0)} \cdot 19^{\min(0,1)}$

$$= 2^2 \cdot 3^0 \cdot 19^0 = 2^2 \cdot 1 \cdot 1$$

i.e. $\text{g.c.d.}(48,76) = 4$

Method 2 (Euclid's Method): Again, say we're asked to find g.c.d. (76,48). First, divide the smaller number into the larger, getting a remainder. $76 = 1 \cdot 48 + 28$. The method of Euclid says: g.c.d. (76,48) = g.c.d. (48,28). Repeating this procedure for g.c.d.(48,28), we first divide 28 into 48, getting $48 = 1 \cdot 28 + 20$. Again Euclid's method says g.c.d. (48,28) = g.c.d.(28,20). Again, dividing 28 by 20, we get $28 = 1 \cdot 20 + 8$. Thus, g.c.d.(28,20) = = g.c.d.(20,8). So, again we divide 20 by 8, giving us $20 = 2 \cdot 8 + 4$. Therefore g.c.d.(20,8) = g.c.d.(8, 4). Now, however, it is easy to see that g.c.d.(8,4) = 4. Thus, but Euclid's method we have that

$$\begin{aligned} \text{g.c.d.}(76,48) &= \text{g.c.d.}(48,28) = \text{g.c.d.}(28,20) = \\ &= \text{g.c.d.}(20,8) = \text{g.c.d.}(8,4) = 4. \end{aligned}$$

So, by the logic of equality, g.c.d.(76,48) = 4.

Lecture 20, May 20, 1968.

We have seen that for any pair of integers x,y there is a g.c.d.(x,y) such that every common divisor of x and y is a divisor of it.

To find the g.c.d.(x,y) we have investigated the two following methods:

Method 1: Obtain prime decomposition of x and y; then use minimum exponents for each prime.

Method 2 (Euclid's Algorithm): Division with remainders. e.g. to find the g.c.d.(136,26), we divide 136 by 26 getting 5 with remainder 6 -- that is, $136 = 5 \cdot 26 + 6$ And Euclid's Algorithm says g.c.d.(136,26) = g.c.d.(26,6) Repeating this process, $26 = 4 \cdot 6 + 2$, and so g.c.d.(26,6) = g.c.d.(6,2), which clearly is 2.

Question: Why does this process of division with remainder give us two numbers whose g.c.d. is the same as the g.c.d. of the original pair of numbers?

In order to answer this question, we will first make clear what we mean by division with remainder and then proceed to a justification of Euclid's Algorithm.

Theorem: Given any positive integers x and y, there are whole numbers q and r such that

$$x = q \cdot y + r \quad \text{and} \quad r < y.$$

We describe these numbers q and r by saying that q is the quotient upon dividing x by y and r is the remainder.

In the above example where we found the g.c.d.(136,26), we first found that $136 = 5 \cdot (26) + 6$. In this example 136 is x, 26 is y, 5 is q and 6 is r.

If we use this theorem, then the key step in Euclid's Algorithm is the

Theorem: $\text{g.c.d.}(x,y) = \text{g.c.d.}(y,r)$

Our proof of this theorem will thus be the justification ~~for the Euclidean Algorithm. We divide the proof into two~~ parts:

Proof Part 1: Suppose z is any common divisor of y and r. Then we claim that z must also be a common divisor of x and y.

Proof of claim: We're assuming z is a common divisor of y and r, so by definition of divisor $y = z \cdot a$ for some integer a, and also $r = z \cdot b$ for some integer b. By our 1st theorem we know $x = q \cdot y + r$. Into this equation we substitute the values of y and r we just found (using the logic of equality).

Thus $x = q \cdot (z \cdot a) + z \cdot b$

Now using the commutative and associative laws for multiplication together with the distributive law, we transform the above equation into $x = z \cdot (q \cdot a + b)$ Since q is an integer (all whole numbers are integers), a is an integer and b is an integer, by the closure law for multiplication and addition we know $q \circ a + b$ is an integer. Hence by definition of divisor, since z multiplied by some integer gives us x, we know z is a divisor of x. Since we assumed z was a divisor of y, we now know z is a common divisor of x and y.

Proof Part 2: Suppose w is any common divisor of x and y. Then we claim w must also be a common divisor of y and r.

Proof of claim: By assumption, $x = w \cdot p$ for some integer p and $y = w \cdot n$ for some integer n. From the theorem about division with remainder, we know $x = y \cdot q + r$. Adding the integer $^{-}(y \cdot q)$ to both sides gives us

$$x - (y \cdot q) = (y \cdot q + r) - y \cdot q ,$$

which reduces to $x - (y \cdot q) = r$.

Now replacing x and y by what we found them equal to gives us $w \cdot p - ((w \cdot n) \cdot q) = r$. Using the necessary general laws (what are they?) gives us $w(p - n \cdot q) = r$ But $p - n \circ q$ is some integer (why?). Thus by definition w must be a divisor of r, and since we were assuming w a divisor of y, we now know w is a common divisor of y and r.

The proofs of Part 1 and Part 2 are now complete.

Putting these two parts together tells us that all the common divisors of x and y are the same common divisors of y and r. So in particular it must be the case that the greatest common divisor of x and y is the greatest common divisor of y and r -- i.e. g.c.d.(x,y) = g.c.d.(y,r).

The Concept of Least Common Multiple

Definition: Given any integers u,v we say that u is a multiple of v just in case (if and only if) v is a factor of u -- i.e., u is a multiple of v just in case $u = a \cdot v$ for some integer a. u is a common multiple of v and t if and only if u is a multiple of v and u is a multiple of t.

Query: Does every pair x,y of integers have a common multiple? Yes, since the integer $x \cdot y$ is a multiple of x and also a multiple of y.

Now there is a theorem which says that for u,v any integers different from 0, there is a smallest positive integer which is a common multiple of u and v. By definit[illegible], we call this integer the least common multiple of u and v and write it as l.c.m.(u,v). Can you see that every common multiple of u and v is a multiple of l.c.m.(u,v)?

How to find the l.c.m. of two given non-zero integers.

We illustrate by two methods, finding l.c.m.$(14,21)$

Method 1: Express both numbers as a product of primes, then find the product of the primes with maximum exponent.

We illustrate: $14 = 2 \cdot 7 = 2^1 \cdot 7^1 = 2^1 \cdot 3^0 \cdot 7^1$

$$21 = 3 \cdot 7 = 3^1 \cdot 7^1 = 2^0 \cdot 3^1 \cdot 7^1$$

So $\text{l.c.m.}(14, 21) = 2^{\max(1,0)}\ 3^{\max(0,1)} \cdot 7^{\max(1,1)}$

$$= 2^1 \cdot 3^1 \cdot 7^1 = 42$$

Method 2: First use Euclid's Algorithm to obtain the g.c.d. of the given pair of numbers. Then use the general law which says

$$\text{l.c.m.}(u,v) = \frac{u \cdot v}{\text{g.c.d.}(u,v)}$$

So again using the two numbers 14 and 21 we first find the g.c.d.(14, 21) by Euclid's Algorithm.

$$21 = 1 \cdot 14 + 7$$

So $\text{g.c.d.}(14, 21) = \text{g.c.d.}(14,7)$

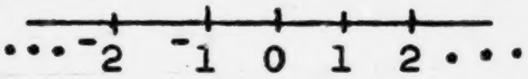
$$= 7$$

Now applying the formula above, we have

$$\text{l.c.m.}(14,21) = \frac{14 \cdot 21}{7} = 42 ,$$

and this checks with method 1.

Our last topic in the system of integers is the relation of order for the integers. Recall that in the system of whole number we had an order relation $<$. When we extend the whole number system of integers we also would like to extend $<$ to a new relation which we'll temporarily write as $<_j$. Also remember that the abstract relation $<$ corresponded to the visual relation "to the left of" when we lined up the whole numbers in their counting order. It is natural that the new relation $<_J$ corresponds to the visual relation "to the left of" on the number line



We make this concept more precise by a

Definition: For x,y any integers we define $x <_J y$ by the following cases:

(i) If x and y are both whole numbers, then we define $x <_J y$ if and only if $x < y$. (This assures us that $<_J$ is an extension of $<$.)

(ii) If x is a negative number and y a whole number then it is always true that $x <_J y$.

(iii) If x a whole number and y a negative number then not $x <_J y$.

(iv) If x and y are both negative numbers then there are natural numbers a,b such that $x = {}^-a$ and $y = {}^-b$ then it will be the case that $x <_J y$ if and only if $b < a$.

Lecture 21, May 22, 1968.

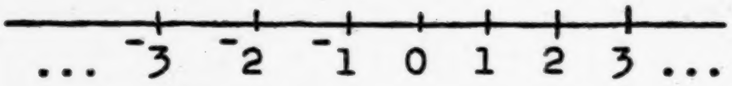
We have extended the whole number system $(W,+,\cdot)$ to ~~a new~~ and larger system, the integer number system $(J,+,\cdot)$. However, this new system also has deficiencies, which we'll now discuss from two points of view, the pure and applied.

I) Pure Mathematics

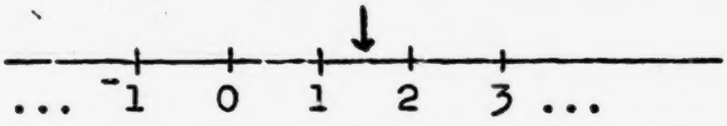
Recall that in the system of whole numbers we can find numbers a and b for which there is no number x such that $a + x = b$. Just take $a = 1$, $b = 0$. But in the system of integers, for every a,b in J we can find an x in J for which $a + x = b$. In fact, x can be taken as $b - a$, for subtraction is always possible in the system $(J,+,\cdot)$, whereas in W we can only form $b - a$ if $b \geq a$. However, in the system J (just as in the system W) we can find numbers a,b with $a \neq 0$ for which there is no number x in J satisfying $a \cdot x = b$. For example, take $a = 2$, $b = 1$. And in general, if a,b are integers with $a \neq 0$, we can find an integer x such that $a \cdot x = b$ only in case $a|b$. That is, a must be a factor of b. The analogy should now be a little clearer, for here we're concentrating on solutions to multiplicative equations, whereas before we wished to find solutions to additive equations. Thus from the standpoint of pure mathematics we wish to extend our number system $(J,+,\cdot)$ to a new number system $(R,+_R,\cdot_R)$ which will satisfy the basic laws studied for earlier systems (e.g., commutative, associative, distributive, cancellation and identity laws) and which in addition will have numbers x satisfying $a \cdot_R x = b$ for any a,b *in R*, as long as $a \neq 0$. Moreover, when we extended $(W,+,\cdot)$ to $(J,+,\cdot)$ we were then able to apply the operation *subtraction* to any a,b in J resulting in a number $b - a$ which is the unique solution to $a + x = b$. Within $(W,+,\cdot)$ a solution to $a + x = b$ was only possible for numbers a,b if $b \geq a$. In the system of integers, for $a \neq 0$ and when $a|b$ we can apply division to obtain an integer $b \div a$, and this is the number x satisfying the equation $a \cdot x = b$. So just as when we moved from W to J, we were then able to apply subtraction to *any* a and b in J; similarly, when we pass from J to R we will be able to apply division to any a,b in R, as long as $a \neq 0$, regardless of whether a is a factor of b or not.

II). Applied Mathematics

a) From the viewpoint of applied mathematics, recall our construction of the number line, an infinite line with integers attached to certain points.



It is clear that there are gaps between adjacent pairs of numbers. Can we find new numbers to attach to the points on the line which are between those points having integers attached to them? Clearly, this would be desirable for purposes of measurement. Additionally, intermediate points on the number line would be of use in helping us to locate points on the plane when we use two perpendicular number lines as a pair of co-ordinate axes. Of interest, also, is the connection between these desired intermediate points on the number line and those numbers x that satisfy $a \cdot x = b,\ a \neq 0$. On the number line consider the point equally distant from the points labelled 0 and 3:



What would be an appropriate number to attach to this point? More precisely if we had such a number, what properties would it have? Well, if we label this point as x and lay off the segment from 0 to x a second time beyond x we reach the point 3. Thus we wish $x + x = 3$. That is,

$$x \cdot 1 + x \cdot 1 = 3,$$

or

$$x \cdot (1 + 1) = 3,$$

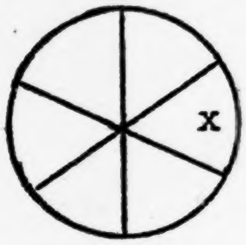
or

$$2 \cdot x = 3,$$

assuming, of course, our new numbers obey the general laws previously studied.

In the same way many numbers attached to intermediate points on the number line can be shown to satisfy multiplicative equations of the form $a \cdot x = b$.

b) A Second Applicational Viewpoint: Pies, the model most often found in elementary school texts. The whole numbers were originally introduced to answer the question "How Many?". i.e. How many objects were in a given set? i.e. Given a set A, what number $n(A)$, do we get by counting the elements of A? In our new extension we would like our new numbers to help us answer the question "How Much?". e.g., given a portion of a whole pie, just how much of the whole pie is represented therein? In other words, we seek a number x to attach to each slice of pie in such a way as to give us a way of "measuring" how much of a whole pie it contains. Once again, there is a connection between these numbers attached to pie slices and those numbers which satisfy $a \cdot x = b$. Suppose we divide a pie equally into six portions:



What can we say about the number x which measures how much of the whole pie is contained in each slice? Clearly, it is a number which when added to itself six times gives us a whole pie. That is, $x + x + x + x + x + x = 1$,

or $$x \cdot (1 + 1 + 1 + 1 + 1 + 1) = 1$$

or $$6 \cdot x = 1$$

Analogously, suppose we want to divide two pies equally among five people. The number y attached to each slice would then clearly satisfy $5 \cdot y = 2$.

c) The following applicational viewpoint, Probability, has not generally been discussed in the elementary school; nevertheless, in the world around us, it is an extremely pervasive concept. The classic example, and one extremely useful, is coin tossing experiments or selection of balls from a box (or urn).

Lecture 22, May 24, 1968.

At the last lecture we saw that an extension of the integer number system $(J,+,\cdot)$ was desireable in order to solve certain multiplicative equations--that is, the integers did not contain numbers that satisfied certain equations and this deficiency has thus provided the motivation for constructing a new number system. However, what exactly should we construct? Our approach to this question can perhaps be clarified by an analogy. If we were asked to go to the forest and seek an animal called a Frumkin, about which we knew nothing, it would be an impossible undertaking; if, however, we found out that this Frumkin has four legs, a brown-ringed tail, fourteen violet whiskers and travels about in packs of no less than seven other Frumkins, our task would certainly be an easier one.

What do we know about what we are seeking? We know we are looking for some number system, we'll call it $(R,+_R,\cdot_R)$, such that:

(i) It is an extension of the previous system $(J,+,\cdot)$.

(ii) The new system sould satisfy the commutative, associative, distributive, cancellation and identity (for 0 and 1) laws. For example we want this new system to satisfy the restricted cancellation law for $\cdot_R$ -- For every x,y,z in R, if $x \cdot_R y = x \cdot_R z$ and if $x \neq 0$, then $y = z$.

(iii) For any numbers a, b in R, if $a \neq 0$ then there is a number x in R such that $a \cdot_R x = b$ (note that this property is not satisfied in $(J, +, \cdot)$)

Let us begin our search for $(R, +_R, \cdot_R)$ as we did when we extended $(W, +, \cdot)$.

Part I: Axiomatic Approach--let us simply assume that we have found a number system $(R, +_R, \cdot_R)$ satisfying properties (i)-(iii) above and explore the consequences of this assumption. This is to help us later on in our definitional approach.

With this assumption, we start with a

Theorem: Given any numbers a, b in R with $a \neq 0$, there is exactly one--no more--number x in R such that $a \cdot_R x = b$

Remark: (iii) above guarantees the fact there will be one number x such that $a \cdot_R x = b$. This theorem tells us there are no more than one.

Proof: Suppose a, b are any numbers in R with $a \neq 0$ and suppose that x and y are numbers in R such that $a \cdot_R x = b$ and also $a \cdot_R y = b$. In order to prove this theorem, we must now show that $x = y$, for then there can be but one solution to $a \cdot_R x = b$. First, by the logic of equality we see that $a \cdot_R x = a \cdot_R y$. Also by assumption we know $a \neq 0$, so we can apply the restricted cancellation law for $\cdot_R$, so $x = y$

This theorem becomes the basis for a

Definition: Given any numbers a, b in R with $a \neq 0$, we define the notation $\frac{b}{a}$ to be the unique number x of R such that $a \cdot_R x = b$.

(How to we know this number x is unique?)

Warning: The textbook uses $\frac{b}{a}$ as a symbol for the ordered pair (b,a).

1st Question: Suppose a,b,c,d are numbers in R with $a \neq 0$, $c \neq 0$. Can we have $\frac{b}{a} = \frac{d}{c}$?

Answer: Of course; whenever $b = d$ and $a = c$.

More Interesting Question: Are there other times? That is, can we have $\frac{b}{a} = \frac{d}{c}$ even if $b \neq d$ and $a \neq c$?

The answer to this 2nd question is yes, as shown by the

Theorem: Let a,b be numbers in R with $a \neq 0$ and let m be any number in R such that $m \neq 0$. Then $\frac{b}{a} = \frac{b \cdot_R m}{a \cdot_R m}$

Proof: By definition, $\frac{b}{a}$ is the unique number x such that

$$\text{(i)} \quad a \cdot_R x = b$$

Also by definition, $\frac{b \cdot_R m}{a \cdot_R m}$ is the unique number y such that

$$\text{(ii)} \quad (a \cdot_R m) \cdot_R y = b \cdot_R m$$

Multiplying both sides of (i) by m, we get

$$(a \cdot_R x) \cdot_R m = b \cdot_R m$$

which we can rearrange by the commutative and associative laws for $\cdot_R$ to give us

$$\text{(iii)} \quad (a \cdot_R m) \cdot_R x = b \cdot_R m$$

Now compare (ii) and (iii). Recall that the y in (ii) was unique. Thus in (iii) x must also be $\frac{b \cdot_R m}{a \cdot_R m}$

Hence $\frac{b}{a} = \frac{b \cdot_R m}{a \cdot_R m}$, and the proof is complete.

Question: Can we have numbers a, b, c, d in R with $a \neq 0$, $c \neq 0$ such that

$$\frac{b}{a} \neq \frac{d}{c} ?$$

We're asking here if there exist fractions in R that are not the same, for perhaps all fractions in R are the same. This is not so and we show this by example. Since R is an extension of J, we know all our old numbers (the integers) must be in R. Using this fact together with the fact that $2 \neq 3$, we show that $\frac{2}{1} \neq \frac{3}{1}$.

Notice that $\frac{2}{1} = 2$ because $\frac{2}{1}$ is the unique x such that $1 \cdot x = 2$. But we know $1 \cdot 2 = 2$. Hence $\frac{2}{1} = 2$. Also $\frac{3}{1} = 3$ since $\frac{3}{1}$ is the unique y such that $1 \cdot y = 3$. But we know $1 \cdot 3 = 3$. Thus $\frac{3}{1} = 3$. Since $2 \neq 3$, we know $\frac{2}{1} \neq \frac{3}{1}$ by the logic of equality. So we see there are in fact different fractions. More generally, we have the

Theorem: If $a \neq 0$, $b \neq d$, then $\frac{b}{a} \neq \frac{d}{a}$

Proof: We leave the proof as an instructive exercise.

Combining the above two theorems on fractions, we get the following criterion for fractions:

$$\frac{b}{a} = \frac{d}{c} \quad \text{if} \quad b \cdot_R c = a \cdot_R d$$

and

$$\frac{b}{a} \neq \frac{d}{c} \quad \text{if} \quad b \cdot_R c \neq a \cdot_R d$$

Lecture 23, May 27, 1968.

We began our axiomatic approach to the rational number system by assuming

(i) $(R, +_R, \cdot_R)$ is a number system, an extension of $(J, +, \cdot)$

(ii) Certain laws hold for $(R, +_R, \cdot_R)$

(iii) For any numbers a, b in R, if $a \neq 0$ then there is some x in R such that $a \cdot_R x = b$.

We proved last time the following theorems:

1. Given a, b in R with $a \neq 0$, there is exactly one x in R satisfying $a \cdot_R x = b$.

   This theorem was the basis for defining $\frac{b}{a}$ to be the unique number x determined by a and b.

2. If a, b are any numbers in R with $a \neq 0$, then $\frac{b}{a} = \frac{b \cdot_R m}{a \cdot_R m}$ for every number m in R that is not zero.

Another theorem (which we did not prove) gave conditions when fractions were not equal.

3. If $a \neq 0$ and $b \neq c$, then $\frac{b}{a} \neq \frac{c}{a}$.

Try to prove this--go back to the definitions of $\frac{b}{a}$ and $\frac{c}{a}$.

Problem: Given two fractions $\frac{b}{a}$ and $\frac{c}{d}$ with $a \neq 0$, $d \neq 0$, how can we decide whether or not $\frac{b}{a} = \frac{c}{d}$?

Solution: Using Theorem 2, we try to find suitable numbers f, g and e (with $e \neq 0$) such that $\frac{b}{a} = \frac{f}{e}$ and $\frac{c}{d} = \frac{g}{e}$. Having done this we know that if $f = g$ then by logic $\frac{f}{e} = \frac{g}{e}$ and hence $\frac{b}{a} = \frac{c}{d}$, and if $f \neq g$, then $\frac{f}{e} \neq \frac{g}{e}$ by Theorem 3;

hence $\frac{b}{a} \neq \frac{c}{d}$. Notice that the crux of this is finding suitable numbers f,g and e in order to express $\frac{b}{a}$ as $\frac{f}{e}$ and $\frac{c}{d}$ as $\frac{g}{e}$.

We give two ways for doing this.

(i) First way. We know $\frac{b}{a} = \frac{b \cdot_R d}{a \cdot_R d}$ by Theorem 3 and the fact that $d \neq 0$. Similarly $\frac{c}{d} = \frac{c \cdot_R a}{d \cdot_R a}$. Thus, since $a \cdot_R d = d \cdot_R a$ by the commutative law for $\cdot_R$ (which we're assuming), we can now complete the comparison. In more familiar language what we've done is find a common denominator for both fractions, whereupon we can compare the numerators.

(ii) Second way. Similar to (i) except that the number e is chosen as the least common multiple of a and d.

We'll now illustrate both of these ways with a numerical example.

Question: Is $\frac{5}{6}$ the same as $\frac{7}{8}$?

by method (i): $\frac{5}{6} = \frac{5 \cdot 8}{6 \cdot 8} = \frac{40}{48}$.

$$\frac{7}{8} = \frac{7 \cdot 6}{8 \cdot 6} = \frac{42}{48}.$$

Now since $42 \neq 40$ we know that $\frac{5}{6} \neq \frac{7}{8}$.

by method (ii): What is the l.c.m.(6,8)? Easy to see that it is 24, by multiplying 6 successively by natural numbers until we reach a number that is a multiple of 8.

Hence, $\frac{5}{6} = \frac{5 \cdot 4}{6 \cdot 4} = \frac{20}{24}$ and

$$\frac{7}{8} = \frac{7 \cdot 3}{8 \cdot 3} = \frac{21}{24}.$$

Since $21 \neq 20$, we know that $\frac{5}{6} \neq \frac{7}{8}$.

We'll now try to make some contact with our textbook. Let's take a fraction, say $\frac{-8}{6}$; by definition this is the unique x such that $6 \cdot_R x = -8$. We also know from Theorem 2 that:

$$\frac{-8}{6} = \frac{-16}{12} = \frac{-24}{18} = \frac{-32}{24} = \cdots$$
$$= \frac{8}{-6} = \frac{16}{-12} = \frac{24}{-18} = \cdots$$

Our textbook considers these fractions as members of a set

$$\{ \frac{-8}{6}, \frac{-16}{12}, \frac{-24}{18}, \ldots \frac{8}{-6}, \frac{16}{-12}, \frac{24}{-18}, \ldots \}$$

By our definition of fractions the number of elements in the above set is 1, since by our criterion all the elements are equal. However, the text gives the notation $\frac{b}{a}$ a different meaning--in the text $\frac{b}{a}$ denotes the ordered pair (b,a); hence, by the text's definition the above set has an infinite number of elements. The book is taking a definitional approach to the rational numbers and has defined

$$\{(-8,6),(-16,12),(-24,18),\ldots,(8,-6),(16,-12),(24,-18)\} = \frac{-8}{6}$$

We, however, are proceeding axiomatically. Later on we shall actually construct the rational numbers.

Problem: Given rational numbers $\frac{b}{a}$ and $\frac{d}{c}$ where $a,c \neq 0$, can we express $\frac{b}{a} +_R \frac{d}{c}$ as some new fraction $\frac{f}{e}$ for suitable numbers f,e $(e \neq 0)$?

The answer, of course, is yes, and we'll show how we can derive the answer within our axiomatic framework.

First, we'll consider the simple case where a and c are equal. Thus, we wish to find $\frac{b}{a} +_R \frac{d}{a}$.

Recall that: $\frac{b}{a}$ is defined as the unique number x in R such that $a \cdot_R x = b$, and

$\frac{d}{a}$ is defined as the unique number y in R such that $a \cdot_R y = d$.

By the logic of equality, we can add these equations getting:

$$(a \cdot_R x) +_R (a \cdot_R y) = b +_R d.$$

Using the distributive law for R, this can be transformed to

$$a \cdot_R (x +_R y) = b +_R d$$

Hence, by Theorem 1 of today's lecture,

$x +_R y$ is the unique number of R which when multiplied by a gives $(b +_R d)$. But by our definition of fraction this unique number is $\frac{b +_R d}{a}$. Thus, $x +_R y$ is the same as $\frac{b +_R d}{a}$ -- that is, $x +_R y = \frac{b +_R d}{a}$. Since $\frac{b}{a} = x$ and $\frac{d}{a} = y$ we have thus shown that

$$\underline{\frac{b}{a} +_R \frac{d}{a} = \frac{b +_R d}{a}}.$$

This last underlined equation tells us that to add two fractions with the same denominator, add their numerators.

Lecture 24, May 29, 1968.

We first reiterate the criterion discussed at the last lecture.

Given $\frac{b}{a}, \frac{d}{c}$ with $a, c \neq 0$, to determine whether or not these are the same rational number find suitable numbers e, f, g so that

$$\frac{b}{a} = \frac{f}{e} \quad \text{and} \quad \frac{d}{c} = \frac{g}{e}.$$

That is, express both fractions with a common denominator. Then $\frac{b}{a} = \frac{d}{c}$ if $f = g$ and $\frac{b}{a} \neq \frac{d}{c}$ if $f \neq g$.

Methods for finding e, f, g:

(i) Take $e = a \cdot_R c$. Then $f = b \cdot_R c$ and $g = a \cdot_R d$.

(ii) Take $e = \text{l.c.m.}(a, c)$; then find appropriate f, g.
For some obscure reason this is the method customarily used in elementary schools.
In actuality, we can take the new deminator e to be *any* common multiple of the denominators a and c and then find suitable numbers f and g.

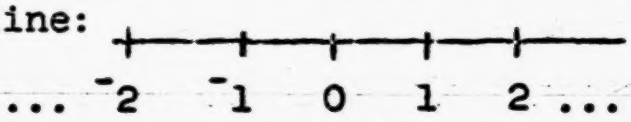
You should have in mind the fact that all our work with fractions has depended on our definition of $\frac{b}{a}$ as the unique number x of R such that $a \cdot_R x = b$. In addition, the proof of the above criterion depended on the following two theorems:

1) Given $\frac{b}{a}$, $a \neq 0$ and given any number $m \neq 0$, then $\frac{b}{a} = \frac{b \cdot_R m}{a \cdot_R m}$. *We proved this in detail.*

2) Given $\frac{f}{e}$ and $\frac{g}{e}$, $e \neq 0$, if $f \neq g$ then $\frac{f}{e} \neq \frac{g}{e}$.

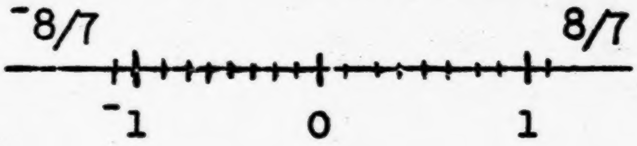
The proof of this theorem was left as an exercise.

Important Remark: In elementary school a theorem such as Theorem 1) above would not be proven. Nevertheless some intuitive evidence for its truth can and should be presented. One way would be to use a number line:



Suppose we're given $\frac{^{-}8}{7}$ and we want to illustrate that $\frac{^{-}8}{7} = \frac{^{-}8 \cdot 2}{7 \cdot 2}$. i.e. $\frac{^{-}8}{7} = \frac{^{-}16}{14}$.

1^{st}: Find that distance which if layed off seven times beginning at the zero point would take us to the point marked 1. Lay this distance off in both directions eight times. Thus we have:



There is a way to now see without laying off any new distances that $\frac{^{-}16}{14} = \frac{^{-}8}{7}$ by taking midpoints of the intervals of $\frac{1}{7}$ distance. Then we'll have a total of 14 equal intervals between 0 and 1 and 16 of these intervals on the negative side of the number line takes us to the point already marked $\frac{^{-}8}{7}$.

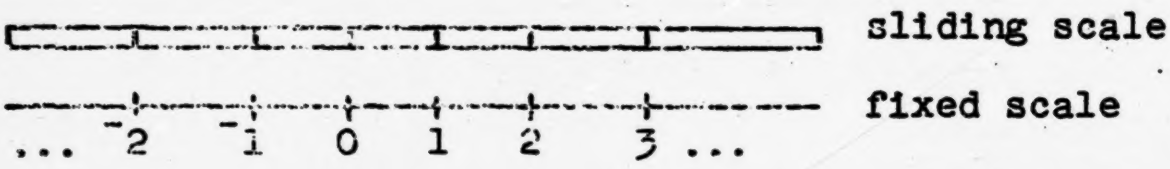
Thus, the arithmetical process of multiplying the numerator and denominator by 2 corresponds to the geometric concept of taking midpoints. This process of dividing the scale into finer intervals is simply a process of using different basic units. Note that any quantity may be expressed in a variety of units; for example, 440 yards is the same as $\frac{1}{4}$ mile. In other words multiplying by a number simply corresponds to a change of scale.

Previously, we discussed the

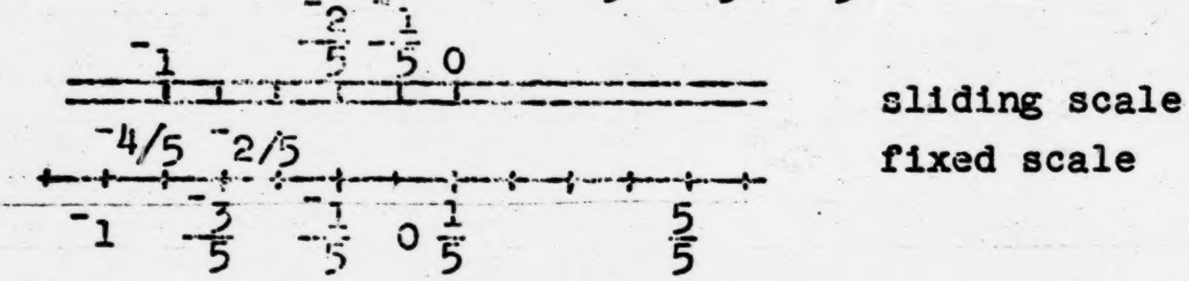
Theorem: Given a, b, c with $a \neq 0$, then

$$\frac{b}{a} +_R \frac{c}{a} = \frac{b +_R c}{a}$$

To illustrate this theorem for the elementary schools, we use a slide rule similar to the one we previously used to add integers like 1 and $^-2$.



To find the sum of fractions like $\frac{1}{5} + \frac{^-2}{5}$, we follow the same basic way as before except we use a change of units, dividing the basic unit of the fixed and sliding scales into fifths. Thus to find $\frac{1}{5} + \frac{^-2}{5}$ we place the 0 of the sliding scale over the $\frac{1}{5}$ point on the fixed scale and read the answer on the fixed scale at the point corresponding to $\frac{^-2}{5}$ on the sliding scale. Thus $\frac{1}{5} + \frac{^-2}{5} = \frac{^-1}{5}$.



To find the sum of fractions with denominators other than 5, the basic unit of the scale would naturally have to be divided differently. One way to do this would be to use strips of clear plastic that would attach to the scales of the slide rule and could be marked on and erased quite easily. These strips might not be part of the regular classroom equipment, and so it would be up to the individual teacher to impress upon the school district her need.

Now, to add two fractions with different denominators, first convert them to fractions with the same denominator.

Example: To add $\frac{b}{a}$ and $\frac{c}{q}$ where $a, q \neq 0$.

Since $\frac{b}{a} = \frac{b \cdot_R q}{a \cdot_R q}$ and $\frac{c}{q} = \frac{a \cdot_R c}{a \cdot_R q}$, and since we know

$$\frac{b \cdot_R q}{a \cdot_R q} +_R \frac{a \cdot_R c}{a \cdot_R q} = \frac{(b \cdot_R q) +_R (a \cdot_R c)}{a \cdot_R q},$$

Thus

$$\frac{b}{a} +_R \frac{c}{q} = \frac{(b \cdot_R q) +_R (a \cdot_R c)}{a \cdot_R q}.$$

Rather than memorize this equation, one can very easily remember that $\frac{b}{a} +_R \frac{c}{a} = \frac{b +_R c}{a}$, using first the method for converting two fractions to fractions with a common denominator.

Recall that we mentioned an application of fractions to the *Theory of Probability*: There are many simple experiments possible for elementary school children which illustrate various laws of probability. For example, there is the law for the probability of the union of two exclusive events: Suppose the probability of event E is a certain rational number $\frac{b}{a}$ and the probability of another event F is $\frac{q}{p}$. Furthermore, suppose that events E and F are exclusive (which means they cannot both happen simultaneously). Then the probability that either event E or event F will occur is $\frac{b}{a} + \frac{q}{p}$. This law may be illustrated as follows: We toss a fair die and designate event E to be: die turns up with upper face showing snake eye (one dot). Since this is a fair die, Probability (E) = $\frac{1}{6}$. We designate as event F: die turns up with an even number of dots on top. Since a die has six sides, three of which have an even number of dots and three of which have an odd number, Probability (F) = $\frac{1}{2}$.

Since events E and F are mutually exclusive (why?)

$$\text{Probability (either E or F)} = \frac{1}{6} + \frac{1}{2} = \frac{(1 \cdot 2) + (1 \cdot 6)}{6 \cdot 2}$$

$$= \frac{2 + 6}{12}$$

$$= \frac{8}{12},$$

or using the l.c.m.(6,2) we'd get

$$\frac{1}{6} + \frac{1}{2} = \frac{1 + 3}{6} = \frac{4}{6}.$$

Lecture 25, May 31, 1968.

In our axiomatic approach to the rational numbers we've assumed $(R,+_R,\cdot_R)$ to be a number system, an extension of $(J,+,\cdot)$, satisfying certain laws and such that for any a,b in R if $a \neq 0$ then there exists $\frac{b}{a} \in R$ with $a \cdot_R (\frac{b}{a}) = b$.

For the operation $+_R$ we obtained the following formulae:

i) $\frac{b}{a} +_R \frac{c}{a} = \frac{b +_R c}{a}$ and

ii) if $m \neq 0$ then $\frac{b}{a} = \frac{b \cdot_R m}{a \cdot_R m}$.

Using these two equations we derived a formula for the sum of fractions of unequal denominator:

If $a,q \neq 0$ then $\frac{b}{a} +_R \frac{c}{q} = \frac{(b \cdot_R q) +_R (c \cdot_R a)}{a \cdot_R q}$

Problem for the operation $\cdot_R$

Given $\frac{b}{a}, \frac{p}{q}$ $(a,q \neq 0)$, can we express the product $\frac{b}{a} \cdot_R \frac{p}{q}$ in the form $\frac{f}{e}$ for suitable numbers f and e $(e \neq 0)$?

Of course, we can. The rule we all learned many years ago is to take $f = b \cdot_R p$ and take $e = a \cdot_R q$. However, we are not interested in simply presenting a formula for multiplication in R. Of more interest is showing how the formula is obtained. This is done in mathematics by a

Proof: by definition of the fractional notation, $\frac{b}{a}$ is the unique number x of R such that

1) $a \cdot_R x = b$ and $\frac{p}{q}$ is the unique number y of R

such that

ii) $q \cdot_R y = p$.

From i) and ii) by the logic of equality we get

iii) $(a \cdot_R x) \cdot_R (q \cdot_R y) = b \cdot_R p$

By several applications of the commutative and associative laws for $\cdot_R$ which are among our axioms (assumptions), iii) becomes

$$(a \cdot_R q) \cdot_R (x \cdot_R y) = b \cdot_R p$$

Using the last equation and what we have been calling Theorem 1, we get that

$x \cdot_R y$ is the unique element of R

which when multiplied by $(a \cdot_R q)$ gives $b \cdot_R p$. But by definition of the fractional notation, this unique element is $\frac{b \cdot_R p}{a \cdot_R q}$. Thus we have shown that $x \cdot_R y = \frac{b \cdot_R p}{a \cdot_R q}$.

Since $x = \frac{b}{a}$ and $y = \frac{p}{q}$ we can conclude that $\frac{b}{a} \cdot_R \frac{p}{q} = \frac{b \cdot_R p}{a \cdot_R q}$ by the logic of equality.

Some applications of multiplication which contribute to an intuitive understanding of it.

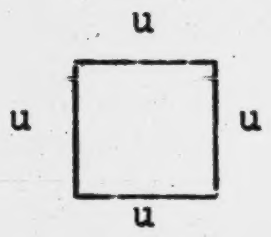
1. Geometry -- the computation of area.

We have all learned that:

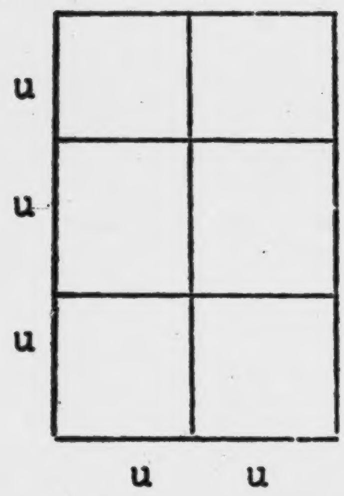
area of a rectangle = (length) $\cdot$ (width).

What does this mean?

Given a unit length u, |———|, this determines a unit area by forming a square all sides of which have the unit u as length:

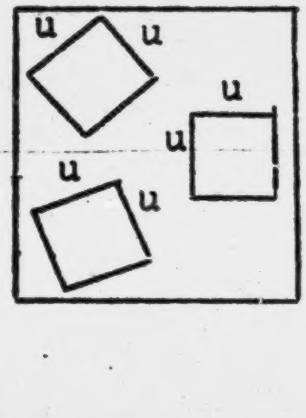


If we have a rectangle and we find the unit length u can be laid off three times on one side and two times for the other side, then using the above formula for area we find that



$$\text{Area} = (3 \cdot 2) \text{ square units}$$
$$= 6 \text{ square units}$$

However, how do we know that if this rectangle is divided up some other way than the above way we'll get the same result for its area? Perhaps starting as we do below and pasting bits and pieces of the unit area here and there on the given rectangle we would come out with 7 square units as the area.

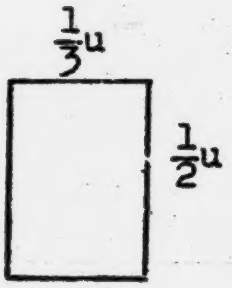


It turns out that 6 square units is indeed the area of this rectangle. We come to believe this by experience, through trial.

Mathematicians, however, are

unwilling to rely on physical methods and have tried and succeeded in proving this within a suitable framework of axioms, but the proof is highly complicated. The theorem which they have proven says that it is impossible to cut up a three by two rectangle using a finite number of cuts and paste it together to form a figure with anything other than six units for area. However, if you are allowed to cut it up into an infinite number of pieces then it is possible. This is called the Banach-Tarski paradox and shows that the notion of area is a deep concept.

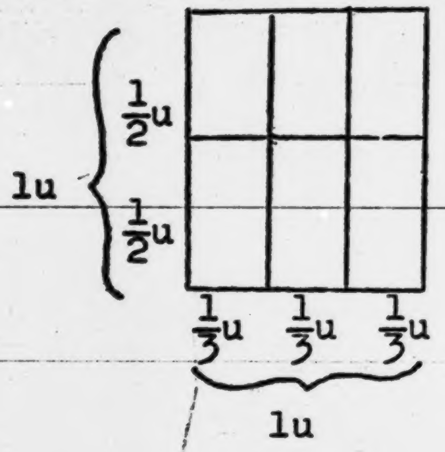
Now suppose we have a rectangle one side of which has length $\frac{1}{2}u$ and the other side has length $\frac{1}{3}u$. The rectangle looks like this:



Applying our formula for area,

$$\text{area of this rectangle} = \left(\frac{1}{2} \cdot_R \frac{1}{3}\right) \text{ square units}$$
$$= \frac{1}{6} \text{ square units.}$$

A way to make this result intuitively convincing for elementary school children would be to look at what we mean by saying the rectangle has a side of length one-half unit. We mean that if we lay that side off twice we get a whole unit of length. Similarly for the length of the other side except we lay it off three times. Thus, if we complete the unit square from these two sides we get the following figure:



Visually one can now see that the unit square is divided into six identical pieces. Thus the original rectangle has as area $\frac{1}{6}$ square units.

Another application for multiplication especially appropriate when dealing with fractions is

2. Probability Theory, and more specifically the computation of the probability of simultaneous occurrences of independent events.

Suppose our "equipment" consists of one fair coin and one fair die.

Let E be the event: die is rolled and comes to rest with 3 dots on its top face.

Let F be the event: coin is rolled and comes to rest with head on top.

Clearly Probability (E) = $\frac{1}{6}$ and Probability (F) = $\frac{1}{2}$

Question: What is the probability of the event that when we roll both the die and the coin the die will rest with

3 dots on top and the coin will rest with the head on top?

According to the theory of probability, as long as the outcome of the coin rolling does not influence the outcome of the die and vice-versa (This is what is meant by independence of the events), then

$$\text{Probability (E and F)} = \frac{1}{6} \cdot \frac{1}{2} = \frac{1}{12}$$

As with area, there are deep questions connected with the theory of Probability. To say that the Probability of the event E above is $\frac{1}{6}$ does not mean that rolling a die 60 times will result in E occurring exactly 10 times. It means that in a vast number of experiments each experiment consisting of rolling the die once, the proportion of times that three dots appear on top would "tend to" $\frac{1}{6}$. This concept of "tends to", however, rests on the notion of limit, a notion studied in that branch of mathematics called analysis.

Lecture 26, June 3, 1968.

Final exam: Friday, June 7, 9-12 220 Hearst Gym.

At our last lecture we introduced multiplication of rational numbers and derived the classical school rule:

$$\frac{b}{a} \cdot_R \frac{d}{c} = \frac{b \cdot_R d}{a \cdot_R c};$$

additionally, we discussed applications to area and Probability Theory.

Question: How can we obtain $\frac{b}{a} \div \frac{d}{c}$ as a rational number of the form $\frac{f}{e}$?

In elementary school the answer that's usually given is a mysterious "Invert and Multiply", meaning

$\frac{b}{a} \div \frac{d}{c}$ is the same as

$\frac{b}{a} \cdot_R \frac{c}{d}$ which is equal to, by our rule for multiplying fractions,

$\frac{b \cdot_R c}{a \cdot_R d}$

Our purpose here, however, is to understand where such cryptic rules come from -- that is, how can it be derived within our axiomatic framework?

First, let's examine what we mean by the notation $\frac{b}{a} \div \frac{d}{c}$.

Definition: We define $\frac{b}{a} \div \frac{d}{c}$ to be the unique z of R such that $z \cdot_R \frac{d}{c} = \frac{b}{a}$.

Since $z \cdot_R \frac{d}{c} = \frac{d}{c} \cdot_R z$ by the commutative law for ${}_R$ (one of our axioms), we have that $\frac{b}{a} \div \frac{d}{c}$ is the unique z

of R such that $\frac{d}{c} \cdot_R z = \frac{b}{a}$. Now from this last equation using our definition of the fractional notation, we get that this unique z is $\frac{\frac{b}{a}}{\frac{d}{c}}$. Hence $\frac{b}{a} \div \frac{d}{c} = \frac{\frac{b}{a}}{\frac{d}{c}}$.

Let us now try to find numbers x and y so that we can express z as $\frac{x}{y}$ -- that is, we're seeking numbers x and y so that $\frac{d}{c} \cdot_R \frac{x}{y} = \frac{b}{a}$. By our rule for multiplying fractions this means finding x and y so that $\frac{d \cdot_R x}{c \cdot_R y} = \frac{b}{a}$. Look carefully at this last equation. Notice that if we could find numbers x and y so that

$$d \cdot_R x = b \quad \text{and} \quad c \cdot_R y = a ,$$

we would then be finished. In the realm of integers, however we know there may not be such numbers x and y. However, if we could change $\frac{b}{a}$ to a new fraction whose numerator is a multiple of d and whose denominator is a multiple of c, then we could solve for x and y in the integers. We can do this by a previous theorem, since

$$\frac{b}{a} = \frac{b \cdot_R (d \cdot_R c)}{a \cdot_R (d \cdot_R c)}$$

Thus, we are now seeking numbers x and y so that

$$\frac{d \cdot_R x}{c \cdot_R y} = \frac{b \cdot_R (d \cdot_R c)}{a \cdot_R (d \cdot_R c)} .$$

Using the commutative and associative laws for $\cdot_R$, we see that we can choose $x = b \cdot_R c$ and $y = a \cdot_R d$. Thus the number z we are looking for is $z = \frac{b \cdot_R c}{a \cdot_R d}$, and hence we have shown that

$$\frac{b}{a} \div \frac{d}{c} = \frac{\frac{b}{a}}{\frac{d}{c}} = \frac{b \cdot_R c}{a \cdot_R d} = \frac{b}{a} \cdot_R \frac{c}{d} .$$

i.e. we have derived the rule for dividing one fraction by another fraction.

This brings to a conclusion our axiomatic approach to the Rational number system. This system can also be reached from the Integers using a definitional approach, whereby we would like to define a set R together with operations $+_R$ and $\cdot_R$ so that:

i) the system $(R, +_R, \cdot_R)$ is an extension of $(J, +, \cdot)$

ii) The basic laws hold for the new system just as they do for the old, and finally the distinguishing property of the new system

iii) Given any a, b in R, if $a \neq 0$ then there is some x in R such that $a \cdot_R x = b$

Actually it is possible to accomplish this in a variety of ways, all of which, hoever, are motivated by first looking at the axiomatic approach.

Our textbook approach, which is the most common definitional approach, is to define the elements of R as sets of the form $\{\ldots(-10,-14),(-5,-7),(5,7),10,14),(15,21),\ldots\}$ or in general $\{(a,b),(m \cdot a, m \cdot b)\}$ where $m = 1,2,3,\ldots,-1,-2,\ldots$ (i.e. m can be any integer except 0) and a and b have no common factor other than 1 and $^{-}1$ and $b > 0$.

Now to define $+_R$ and $\cdot_R$

Given any two such elements of R, say $\{(m\cdot a, m\cdot b)\}$ and $\{(m\cdot c, m\cdot d)\}$ we define $\cdot_R$ to be the operation such that $\{(m\cdot a, m\cdot b)\} \cdot_R \{(m\cdot c, m\cdot d)\} = \{(m\cdot x, m\cdot y)\}$, where x is obtained from $a\cdot c$ and y is obtained from $b\cdot d$ by division by the g.c.d.$(a\cdot c, b\cdot d)$. By looking carefully at our axiomatic approach we analogously define $+_R$.

These above definitions do not exactly give us an extension of $(J,+,\cdot)$ since our old numbers do not actually appear in this new system. However if we replace those elements of R of the form $\{(m\cdot a, m\cdot b)\}$ by the element a of J, we would then have an extension which agrees with what we defined an extension to be.